



OPEN

## Screening obstructive sleep apnea patients via deep learning of knowledge distillation in the lateral cephalogram

Min-Jung Kim<sup>1,6</sup>, Jiheon Jeong<sup>1,2,6</sup>, Jung-Wook Lee<sup>1,6</sup>, In-Hwan Kim<sup>2,6</sup>, Jae-Woo Park<sup>2,6</sup>, Jae-Yon Roh<sup>3</sup>, Namkug Kim<sup>1,4</sup>✉ & Su-Jung Kim<sup>2,5</sup>✉

The lateral cephalogram in orthodontics is a valuable screening tool on undetected obstructive sleep apnea (OSA), which can lead to consequences of severe systematic disease. We hypothesized that a deep learning-based classifier might be able to differentiate OSA as anatomical features in lateral cephalogram. Moreover, since the imaging devices used by each hospital could be different, there is a need to overcome modality difference of radiography. Therefore, we proposed a deep learning model with knowledge distillation to classify patients into OSA and non-OSA groups using the lateral cephalogram and to overcome modality differences simultaneously. Lateral cephalograms of 500 OSA patients and 498 non-OSA patients from two different devices were included. ResNet-50 and ResNet-50 with a feature-based knowledge distillation models were trained and their performances of classification were compared. Through the knowledge distillation, area under receiver operating characteristic curve analysis and gradient-weighted class activation mapping of knowledge distillation model exhibits high performance without being deceived by features caused by modality differences. By checking the probability values predicting OSA, an improvement in overcoming the modality differences was observed, which could be applied in the actual clinical situation.

Obstructive sleep apnea (OSA) is a complex and heterogeneous disorder that draws attention worldwide, because undetected OSA can lead to consequences of severe systematic disease<sup>1,2</sup>, heart disease<sup>3</sup>, cardiovascular dysfunction<sup>4</sup>, stroke<sup>5</sup>, and even sudden death<sup>6</sup>. Therefore, diagnosing OSA as early as possible takes precedence for the proper management of OSA<sup>7</sup>. Polysomnography (PSG) is regarded as the gold standard method used for OSA diagnosis<sup>8</sup>. Nevertheless, PSG is an overnight examination that is expensive and requires high patient compliance; additionally, it has a high risk of invalid study results<sup>9,10</sup>.

One of the major phenotypic causes of OSA is craniofacial structural abnormality<sup>11,12</sup>. The predisposing characteristics of OSA on the craniofacial structures are a narrowed posterior airway space, a long, elongated pharynx, thicker soft palate, a long and large tongue, a lower position of the hyoid bone, increased anterior lower facial height, and decreased sagittal dimension of the cranial base. The lateral cephalogram has been acknowledged as the tool confirming the potential relevance of OSA in patients with suspected symptoms<sup>13–15</sup>.

Deep learning is a subset of artificial intelligence (AI) technique that can learn from special features and make predictions about image data with or without supervision<sup>11</sup>. Amongst the various deep learning approaches, convolutional neural networks (CNNs) have been highlighted in image recognition to recognize anatomical structures in medical images automatically<sup>12</sup>. Savoldi et al.<sup>13</sup> applied lateral cephalograms to evaluate the anatomical structures associated with OSA in children. They reported a limited reliability in the assessment of the tongue and soft palate area using lateral cephalograms. Tsuiki et al.<sup>14</sup> included groups of patients with severe

<sup>1</sup>Department of Convergence Medicine, Asan Medical Institute of Convergence Science and Technology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-Gil Songpa-gu, Seoul 05505, Republic of Korea. <sup>2</sup>Department of Biomedical Engineering, Asan Medical Institute of Convergence Science and Technology, Asan Medical Center, College of Medicine, University of Ulsan, Seoul, Republic of Korea. <sup>3</sup>Department of Orthodontics, Kyung Hee University Dental Hospital, Seoul 05505, Republic of Korea. <sup>4</sup>Department of Radiology, University of Ulsan College of Medicine, Asan Medical Center, 88 Olympic-Ro 43-Gil Songpa-Gu, Seoul 05505, Republic of Korea. <sup>5</sup>Department of Orthodontics, School of Dentistry, Kyung Hee University, 23, Kyunghedae-ro, Dongdaemun-gu, Seoul 02447, Republic of Korea. <sup>6</sup>These authors contributed equally: Min-Jung Kim, Jung-Wook Lee, Jiheon Jeong, In-Hwan Kim and Jae-Woo Park. ✉email: namkugkim@gmail.com; ksj113@khu.ac.kr

OSA and non-OSA according to the craniofacial morphology. They employed the deep learning approach to perform lateral cephalogram-based image classification, with successful identification of individuals with severe OSA by deep CNNs using lateral cephalogram.

Various modalities in cephalograms have become vital in the actual clinical setting<sup>15</sup>. Therefore, robust training in multimodal imaging has drawn attention in the computer science field<sup>16</sup>. Some previous studies have attempted deep learning algorithms with knowledge distillation to overcome the modality differences in images<sup>17,18</sup>. Knowledge distillation was introduced by Geoffrey Hinton et al. in Google Inc<sup>19</sup>. Distilling the knowledge is to compress the knowledge into a single model. This comprises first training the cumbersome model called “teacher model” and subsequently using another type of training to transfer knowledge from the cumbersome model to a small model called the “student model”. Since deep learning is outstanding for training multiple levels of feature representation, feature-based knowledge distillation used both the output of the last layer and the output of feature maps in intermediate layers as the knowledge to supervise the student model training<sup>20</sup>.

Based on that craniofacial phenotypes of OSA patients have observable anatomical intrinsic features in the lateral cephalogram, we attempted to develop a fully automated cephalometric screening tool of OSA presence in a simple manner. We aimed to propose a reinforced CNNs algorithm to perform an OSA-classification task in different cephalometric modalities, and to detect the region of interests (ROIs) which indicate the anatomical risk areas contributing to OSA.

## Materials and methods

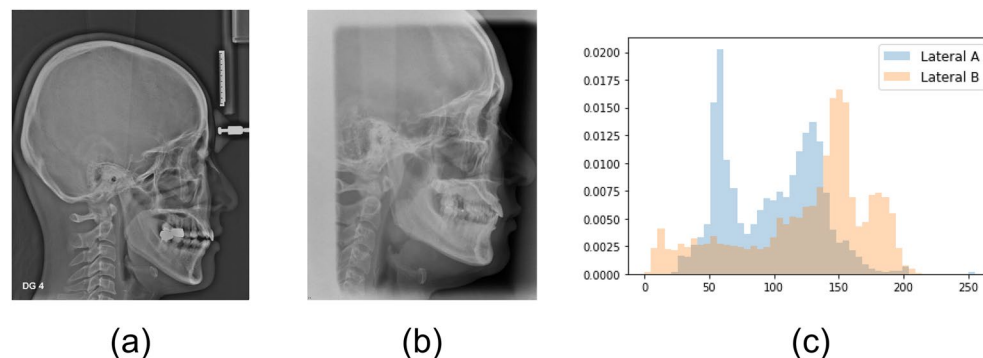
### Data acquisition with lateral cephalogram protocol

Five hundred lateral cephalograms of adult OSA patients, who had been diagnosed by apnea-hypopnea index (AHI) from PSG record, were randomly collected from picture archiving and communication system (PACS) in Kyung Hee University Medical Center and Dental Hospital. This study samples comprised 100 images (Lateral A) taken with CX-90SP machine (Asahi Roentgen, Kyoto, Japan) and 400 images (Lateral B) taken with DP80P machine (Dentsply Sirona, Bensheim, Germany). The lateral cephalograms were taken in a strictly standardized head and jaw postures at the end of expiration during the respiratory cycle to obtain the most relaxed pharyngeal soft tissue images. The OSA group were subdivided into mild ( $5 < \text{AHI} \leq 15$ ,  $n = 100$ ), moderate ( $15 < \text{AHI} \leq 30$ ,  $n = 154$ ), and severe ( $\text{AHI} > 30$ ,  $n = 246$ ) groups according to the disease severity<sup>21</sup>.

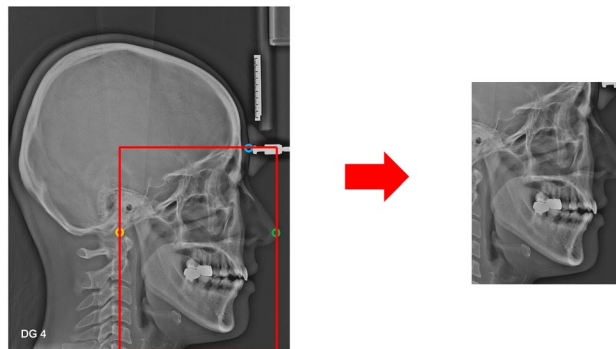
For the control group (non-OSA group), 498 cephalograms of healthy orthodontic patients taken with CX-90SP machine (Lateral A) were recruited. Non-OSA samples were defined as the patients without any OSA signs or symptoms in the clinical examination using sleep questionnaire, instead of AHI number. Finally, a total of 998 cephalograms were included in the present study. All patients provided with informed consents for the anonymous use of their cephalograms and PSG. This retrospective study was performed under the institutional review board for the protection of human subject (IRB Number: KH-DT19006).

### Image pre-processing

The modality differences between Lateral A and Lateral B images existed in the image magnification and resolution as well as the standardization format of image acquisition, which were contrasted by a histogram (Fig. 1). To unify the field of view (FOV) of two image modalities, all images were cropped to include all the potential critical regions of interests (ROIs) predisposing to OSA. Three cephalometric landmarks were automatically extracted with the previously developed landmark-detection algorithm to generate the consistent cropping boundaries: Basion, Glabella, and Pronasale points (Fig. 2)<sup>22</sup>. The accuracy of each landmark selected by algorithm is  $0.76 \pm 1.06$ ,  $1.36 \pm 1.08$ , and  $0.60 \pm 0.46$  mm, respectively<sup>22</sup>. The landmark position selected by algorithm was not modified for cropping. In addition, when resizing to the input shape, the aspect ratio of the original image was maintained to minimize the distortion of the actual image information. For data augmentation, image rotation was applied in a range of  $-15^\circ$  to  $15^\circ$ , and horizontal flip, CLAHE (Contrast Limited Adaptive Histogram Equalization) were applied randomly with a probability of 50%.



**Figure 1.** The modality differences between the two lateral cephalograms: (a) Lateral A image, (b) Lateral B image, and (c) histogram distribution of the two images.



**Figure 2.** Image cropped based on Landmark. (Yellow: “Basion”, Blue: “Glabella”, Green: “Pronasale”).

### Knowledge distillation

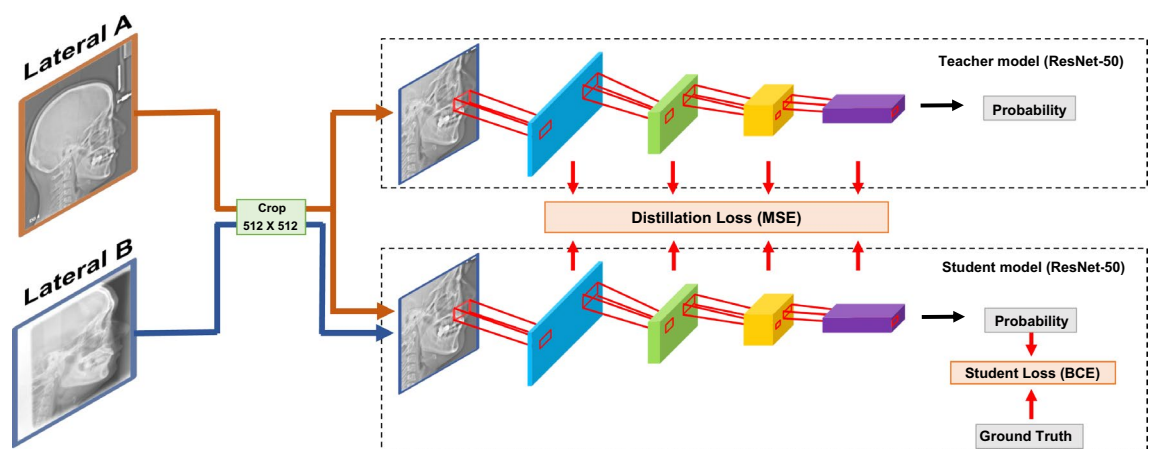
Since an image of each modality exists predominantly for each group, the regular method is highly likely to create a model that classifies by recognizing the difference in the modality rather than classifying based on rational grounds. Therefore, to overcome this problem, we used a feature-based knowledge distillation deep learning model. Simultaneously, we used ResNet-50 to train the data in order to compare with the proposed model. By applying this method, we (1) first extracted only the modality images common to both groups to create a teacher model, and (2) built a model architecture that ultimately created a student model that additionally learns other modality images while maintaining the performance of this teacher model (Fig. 3). Both models used the ResNet-50 architecture as a backbone, and used pre-trained weights with the ImageNet dataset. Moreover, a comparative experiment was conducted with additional information such as age, and BMI. After global average pooling (GAP) of the ResNet-50 architecture, sex encoded to one-hot vector was added and the combined information passed through the fully connected layers to derive the final OSA prediction values.

### Model training

333 (67%) OSA patient images and 333 (67%) non-OSA patient images were used for training; additionally, 84 (17%) images each from OSA and non-OSA patients were used for validation. For testing, 83 (17%) OSA patient images and 81 (16%) non-OSA patient images were used.

### Evaluation

The performances of the deep neural network models were analyzed for accuracy, area under receiver operating characteristic curve (AUROC), sensitivity, and specificity in classification using Python scikit-learn (version 0.24.2). Unpaired *t* tests were used to compare baseline characteristics between the OSA and non-OSA groups. In addition, paired *t* tests were also performed to compare the results of the model with and without sex. Statistical analysis was performed by Python SciPy (version 1.7.1) in the present study with a two-sided *P* value < 0.05 defined as statistically significant. To verify and visualize the critical ROIs that the proposed deep learning model has seen in prediction the OSA patients from the lateral cephalograms, we performed gradient-weighted class activation mapping (Grad-CAM).



**Figure 3.** Diagrams of the model architecture.

## Results

### Unpaired *t* test between OSA and non-OSA groups

As shown Table 1, the OSA group revealed significantly greater age and higher body mass index (BMI) than non-OSA group. The mean AHI of OSA group was  $34.4 \pm 22.2$  events/h. Unpaired *t* test of the OSA and the non-OSA groups showed significant differences in sex ( $p < 0.01$ ), age ( $p < 0.01$ ), and BMI ( $p < 0.01$ ) between the two groups.

### Analysis through OSA predicted probability values (Fig. 4)

In order to check whether the problem of classification by the features appearing in the difference in modality was effectively solved, the OSA prediction probability values of the ResNet-50 and our knowledge distillation models were checked. As shown Fig. 4, the y-axis is the predicted probability value of OSA and the x-axis corresponds to the index number of the data. And the red dot represents the Lateral A image of the non-OSA group, the blue dot is the Lateral A image of the OSA group, and the green dot is the Lateral B image of the OSA group. Through the results of ResNet-50 model (Fig. 4a), it can be confirmed that the red and green dots are predicted with strong certainty as non-OSA and OSA, respectively. On the other hand, the blue dots are evenly distributed over the total probability values. Looking at the result of our model (Fig. 4b), it was observed that red and green dots are generally found to be located on the non-OSA and OSA side, respectively, as they are more evenly distributed than in ResNet-50 model (Fig. 4a), which shows an extreme pattern. Moreover, the blue dot is mostly located on the OSA side.

### Accuracy, AUROC, sensitivity, and specificity of the test set

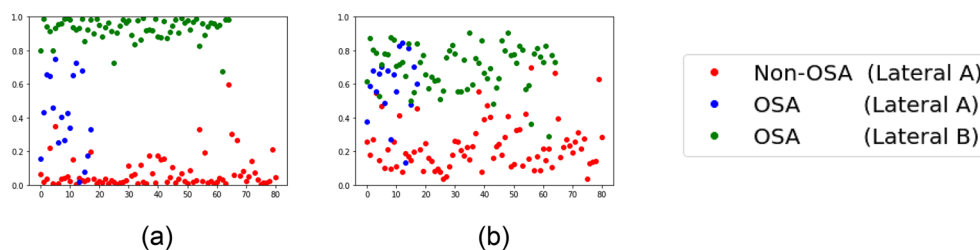
With respect to the result of the entire dataset, the basic model had the highest values of accuracy, AUROC, and specificity (0.927, 0.982, and 0.988) as shown Table 2. On the other hand, sensitivity had the highest value when cropped in the student model (0.940). In addition, the student model had higher values of accuracy, AUROC, and sensitivity than the teacher model although the specificity was slightly lower.

### Analysis of results of our model according to OSA severity

To compare the results according to the OSA severity in our model, we checked the performance of each OSA group in the test set. The severe group ( $n = 40$ ) showed the highest AUROC (0.990), followed by the mild group ( $n = 18$ ) and the moderate group ( $n = 25$ ) (Table 3 and Fig. 5). The results of confusion matrix in each OSA group are shown Fig. 6, respectively.

Patient characteristics	OSA	non-OSA
n	500	498
Age (years)	$45.3 \pm 12.5$	$27.2 \pm 8.9$
BMI ( $\text{kg}/\text{m}^2$ )	$25.6 \pm 3.2$	$22.2 \pm 4.0$
AHI (events /h sleep)	$34.4 \pm 22.2$	
Male:female	400:100	205:293

**Table 1.** Baseline demographics of the two groups (OSA, Non-OSA).



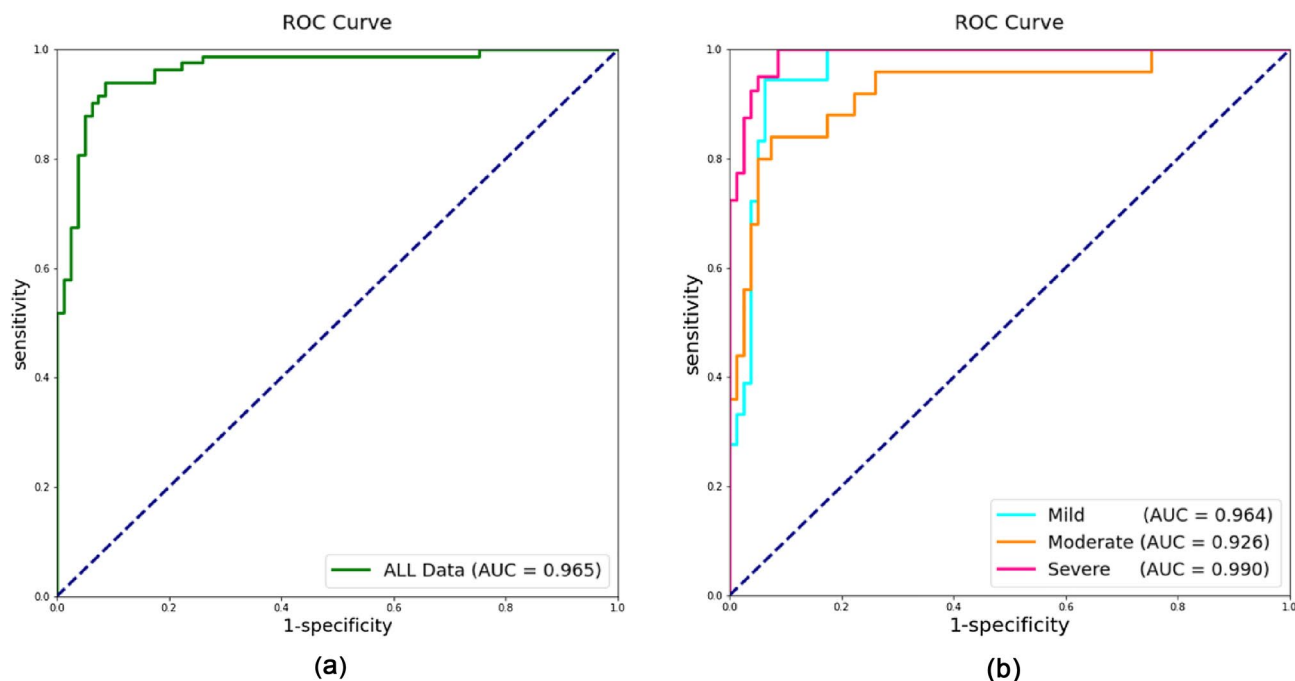
**Figure 4.** Probability predicted as OSA according to the modality. (a) ResNet-50 Model, (b) Our Model.

	Accuracy	AUROC	Sensitivity	Specificity
Base model	0.927	0.982	0.867	0.988
Teacher model	0.841	0.932	0.771	0.914
Student model	0.921	0.965	0.940	0.901

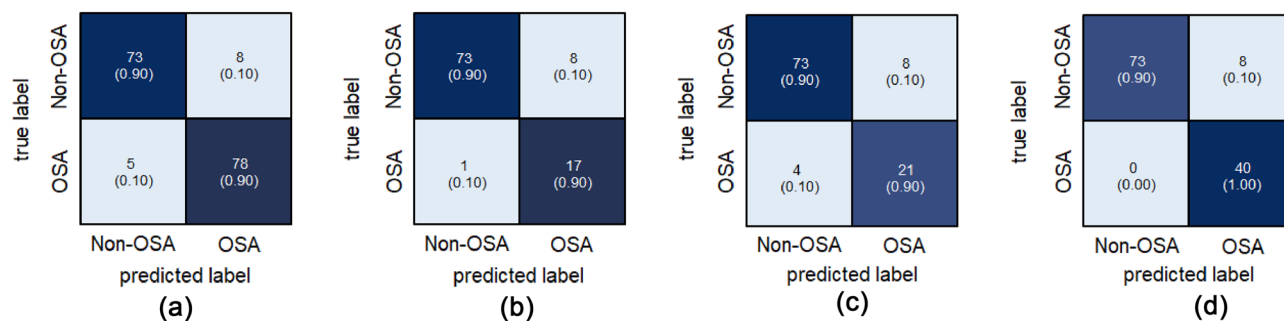
**Table 2.** The accuracy, AUROC, sensitivity, and specificity of the test set in the binary receiver operating characteristic (ROC) analysis according to the model type.

Severity	Accuracy	AUROC	Sensitivity	Specificity
Our model				
Mild	0.909	0.964	0.944	0.901
Moderate	0.887	0.926	0.840	0.901
Severe	0.934	0.990	1.000	0.901

**Table 3.** Predicted values of our model according to severity of OSA.



**Figure 5.** Receiver operating characteristics curves of our model to classify the OSA (a) Classifying the OSA and non-OSA. (b) Analysis result of OSA according to the severity.



**Figure 6.** Confusion matrix results of binary classification according to the OSA severity (a) all OSA, (b) mild, (c) moderate and (d) severe.

### Grad-CAM for diagnosis OSA

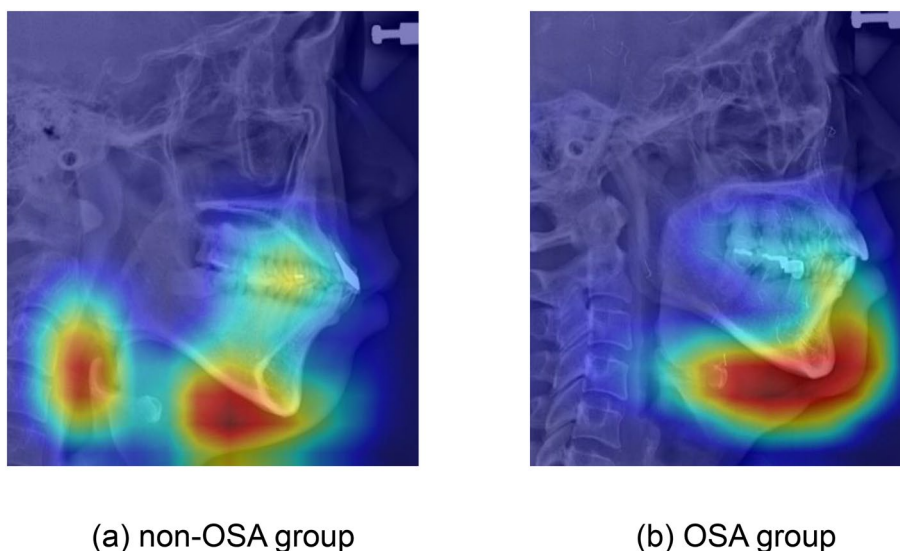
To confirm the results of Grad-CAM in a more objective and universal manner, the results were derived as the average value of all the test sets Table 4. Before applying the average value, registration was applied to all datasets based on landmarks. The results showed different patterns between the OSA and non-OSA groups in Fig. 7. First, in the case of the non-OSA group, our model tends to classify by focusing on the airway and the submental area. On the other hand, our model focused on a wider area, including not only the submental, but also the tip of the chin.

### Comparison of the performance of a classification model with additional information

To confirm the performance change of the model when additional information is included, the performance of the entire test set was compared. The accuracy was higher when learning using only sex than when learning

	Accuracy	AUROC	Sensitivity	Specificity
Teacher model	0.841	0.932	0.771	0.914
Student model+ sex	0.909	0.962	0.952	0.864
Student model+ sex + BMI	0.804	0.792	0.682	0.830
Student model+ sex + age	0.822	0.816	0.712	0.860
Student model+ sex + BMI + age	0.839	0.828	0.721	0.881

**Table 4.** Comparison of the accuracy, AUROC, sensitivity, specificity of the test set between the models with additional information.



**Figure 7.** Average result of test set Grad-CAM (a) non-OSA group and (b) OSA group.

BMI or age together. This might be due to differences in gender differences between the normal group and the OSA group. When Age and BMI were additionally learned, the accuracy was similar to that of the teacher model.

## Discussion

In our study, we proposed a knowledge distillation-based deep learning model to classify the patients into the OSA and non-OSA groups using the lateral cephalogram. Lateral cephalograms have been used as a tool to evaluate the craniofacial morphology in patients with malocclusion in the Department of Orthodontics. Since patients with OSA exhibit anatomical features, which are observable in lateral cephalograms, it is useful to examine whether deep learning can be a screening tool for lateral cephalograms to classify OSA and non-OSA. Tsuiki et al.<sup>14</sup> attempted to develop a deep learning-based OSA screening tool using the lateral cephalogram for the first time. However, their study did not analyse all OSA severity and included only patients with severe OSA. In the real world, since the imaging device used by each hospital is different, the inevitable differences of image modality should be considered. Therefore, we tried to overcome this problem and developed a robust deep learning model. During the first training using ResNet-50, even though all the images were processed to overcome modality differences, the problem of seeing and classifying OSA with modality still existed because most lateral cephalograms of OSA patients were scanned by a specific imaging modality, DP80P, from the Department of Otolaryngology.

The non-OSA group includes all Lateral A images, whereas the OSA group includes mostly Lateral B images; however, there are also a few Lateral A images. Therefore, by checking the result value of the Lateral A image among the OSA groups, it is possible to evaluate whether the problem caused by the difference in modality has been overcome (corresponding to the blue dot in Fig. 3). Based on the results of the base ResNet-50 model, in the case of blue dots, it was confirmed that they were spread throughout the entire probability values and were particularly biased toward non-OSA. This can be interpreted as the result of the model being classified by recognizing only the same modality information, since the majority of the “Lateral A” corresponds to non-OSA. In contrast, in our model, most of the blue dots are in the OSA group and the green and red dots are also not extremely biased to one side, but rather spread more naturally at a classifiable level; therefore, it can be said that learning was accomplished by overcoming the differences in the modality.

Compared to the model of Tsuiki et al.<sup>14</sup>, our model has slightly higher scores for AUROC, sensitivity, and specificity (Table 3). Although our data are small in number and the difference in modality has significant disadvantages, knowledge distillation overcomes these problems<sup>23</sup>, and the final result is stable and excellent

performance is not biased against a specific class compared to the previous study ( $n = 1389$  lateral cephalogram images).

In order to make a more accurate and rational classification model, it is right to learn by adding clinical information other than the images necessary for diagnosing OSA. Therefore, we trained a model with sex information; however, the difference in the performance was insignificant. This is because, first, there is a difference in sex between the two groups, and information about sex could be recognized by itself only through image during training. Therefore, in providing additional information, it is important to select which information to choose, and it is necessary to think about how to provide that information harmoniously with the information of the image.

When discussing the OSA potential, there were four major OSA-related anatomical characteristics that often pay attention which are narrowed pharynx, submental, chin, and nasomaxillary complex. Considering average result of Grad-CAM in Fig. 7, our model focuses on the area around the tip of the chin and the submental area both in the non-OSA and OSA groups. It means that our AI model classified OSA in a way similar to what humans see attentively. In addition, the airway is also an important area in diagnosing OSA. However, unlike the non-OSA group, it is not an area of interest in the OSA group. Normal anatomical features with wide airways can be recognized and classified as normal; however, irregular features of the OSA group, such as narrow or those blurred by the edges of the bones, are considered to be relatively unnoticed in the classification.

There are several limitations in this study. According to Table 1 and the results of the unpaired  $t$  test, it can be seen that the BMI of the OSA group and the non-OSA group are significantly different, which could lead to classifying by focusing only on the degree of obesity. In fact, it was confirmed that Grad-CAM also focused more on the submandibular fat layer. Moreover, unlike the study by Tsuiki et al.<sup>14</sup>, where only the severe OSA group showing the characteristics of a more crowded oropharynx was tested, we included all the OSA groups in the experiment. Therefore, data with ambiguous characteristics may exist and non-anatomical OSA patients may exist, so it is thought that it was difficult to create an ideal model that can classify by considering all four areas (narrowed pharynx, submental, chin, and nasomaxillary complex) as the OSA areas of interest. For further study, additional clinical information should be considered to train a model with larger sample size. Second, there is no external validation in this study. It is very difficult to acquire an additional external dataset. In the near future, this model should be tested for external validation for evaluation of overfitting and robustness of this model.

## Conclusion

A suitable deep CNN successfully classified the patients with OSA and non-OSA using a 2-dimensional lateral cephalogram. These excellent results were obtained regardless of the severity of the OSA. Even in cephalograms obtained from different X-ray modalities, OSA patients could be accurately identified through appropriate knowledge distillation.

## Data availability

This retrospective study was conducted according to the principles of the Declaration of Helsinki, and was performed in accordance with current scientific guidelines. The study protocol was approved by the Institutional Review Board Committee of Kyung Hee University Medical Center and Dental Hospital, Korea. (KH-DT19006). The requirement for informed patient consent was waived by the Institutional Review Board Committee of Kyung Hee University Medical Center and Dental Hospital.

Received: 21 November 2022; Accepted: 15 September 2023

Published online: 18 October 2023

## References

- Ban, W. H. & Lee, S. H. Obstructive sleep apnea and chronic lung disease. *Chronobiol. Med.* **1**(1), 9–13. <https://doi.org/10.33069/cim.2019.0001> (2019).
- Botros, N. *et al.* Obstructive sleep apnea as a risk factor for type 2 diabetes. *Am. J. Med.* **122**(12), 1122–1127. <https://doi.org/10.1016/j.amjmed.2009.04.026> (2009).
- Parati, G. *et al.* Heart failure and sleep disorders. *Nat. Rev. Cardiol.* **13**(7), 389–403. <https://doi.org/10.1038/nrcardio.2016.71> (2016).
- Somers, V. K. *et al.* Sleep apnea and cardiovascular disease: an American heart association/American college of cardiology foundation scientific statement from the American heart association council for high blood pressure research professional education committee, council on clinical cardiology, stroke council, and council on cardiovascular nursing. *J. Am. Coll. Cardiol.* **52**(8), 686–717. <https://doi.org/10.1016/j.jacc.2008.05.002> (2008).
- Sharma, S. & Culebras, A. Sleep apnoea and stroke. *Stroke Vasc. Neurol.* **1**(4), 185–191. <https://doi.org/10.1136/svn-2016-000038> (2016).
- Chan, A. & Antonio, N. Mechanism of sudden cardiac death in obstructive sleep apnea, revisited. *Sleep Med.* **14**, e95. <https://doi.org/10.1016/j.sleep.2013.11.201> (2013).
- Davis, A. P., Billings, M. E., Longstreth, W. T. Jr. & Khot, S. P. Early diagnosis and treatment of obstructive sleep apnea after stroke: Are we neglecting a modifiable stroke risk factor?. *Neurol. Clin. Pract.* **3**(3), 192–201. <https://doi.org/10.1212/CPJ.0b013e318296f274> (2013).
- Tan, H.-L., Gozal, D., Ramirez, H. M., Bandler, H. P. R. & Kheirandish-Gozal, L. Overnight polysomnography versus respiratory polygraphy in the diagnosis of pediatric obstructive sleep apnea. *Sleep* **37**(2), 255–260. <https://doi.org/10.5664/sleep.3392> (2014).
- Pang, K. P. & Terris, D. J. Screening for obstructive sleep apnea: An evidence-based analysis. *Am. J. Otolaryngol.* **27**(2), 112–118. <https://doi.org/10.1016/j.amjoto.2005.09.002> (2006).
- Kapur, V. K. *et al.* Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: An American academy of sleep medicine clinical practice guideline. *J. Clin. Sleep Med.* **13**(3), 479–504. <https://doi.org/10.5664/jcs.6506> (2017).
- Haskins, G., Kruger, U. & Yan, P. Deep learning in medical image registration: A survey. *Mach. Vis. Appl.* **31**(1), 8. <https://doi.org/10.1007/s00138-020-01060-x> (2020).
- Shen, D., Wu, G. & Suk, H.-I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* **19**(1), 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442> (2017).

13. Savoldi, F. *et al.* Reliability of lateral cephalometric radiographs in the assessment of the upper airway in children: A retrospective study. *Angle Orthod.* **90**(1), 47–55. <https://doi.org/10.2319/022119-131.1> (2019).
14. Tsuiiki, S. *et al.* Machine learning for image-based detection of patients with obstructive sleep apnea: An exploratory study. *Sleep Breath.* <https://doi.org/10.1007/s11325-021-02301-7> (2021).
15. Guo, Z., Li, X., Huang, H., Guo, N. & Li, Q. Deep learning-based image segmentation on multimodal medical imaging. *IEEE Trans. Radiat. Plasma Med. Sci.* **3**(2), 162–169. <https://doi.org/10.1109/TRPMS.2018.2890359> (2019).
16. Liu, W. & Tao, D. Multiview hessian regularization for image annotation. *IEEE Trans. Image Process.* **22**(7), 2676–2687. <https://doi.org/10.1109/TIP.2013.2255302> (2013).
17. Liu, Y., Wang, K., Li, G. & Lin, L. Semantics-aware adaptive knowledge distillation for sensor-to-vision action recognition. *IEEE Trans. Image Process.* **30**, 5573–5588. <https://doi.org/10.1109/TIP.2021.3086590> (2021).
18. Thoker, F. M & Gall, J. Cross-modal knowledge distillation for action recognition 6–10 (2019).
19. Geoffrey Hinton, O. V. & Jeff, D. Distilling the knowledge in a neural network (2015).
20. Gou, J., Yu, B., Maybank, S. J. & Tao, D. Knowledge distillation: A survey. *Int. J. Comput. Vis.* **129**(6), 1789–1819. <https://doi.org/10.1007/s11263-021-01453-z> (2021).
21. Young, T., Peppard, P. E. & Gottlieb, D. J. Epidemiology of obstructive sleep apnea. *Am. J. Respir. Crit. Care Med.* **165**(9), 1217–1239. <https://doi.org/10.1164/rccm.2109080> (2002).
22. Kim, I.-H., Kim, Y.-G., Kim, S., Park, J.-W. & Kim, N. Comparing intra-observer variation and external variations of a fully automated cephalometric analysis with a cascade convolutional neural net. *Sci. Rep.* **11**, 7925 (2021).
23. Park, W., Kim, D., Lu, Y. & Cho, M. Relational knowledge distillation 3967–3976 (2019).

## Acknowledgements

Special thanks to S.J.K and J.Y.R for their help in establishing a dataset for the deep learning. This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI18C2383, HI21C1148).

## Author contributions

M.J.K., J.W.L. and J.H.J contributed to data acquisition, deep learning model development, interpretation of data, and draft and critical revision of the manuscript. S.J.K. contributed to conception, design, data acquisition, analysis, interpretation, draft, and critical revision of the manuscript. N.K. contributed to conception, design, and critical revision of the manuscript. All authors gave final approval and agreed to be accountable for all aspects of the work.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to N.K. or S.-J.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023