A Deep Learning Approach to Localizing Multi-level Airway Collapse Based on Snoring Sounds

Ying-Chieh Hsu[1,2,3,*], Stanley Yung-Chuan Liu[4], Chao-Jung Huang[4,5], Chi-Wei Wu[6], Ren-Kai Cheng[6], Jane Yung-Jen Hsu[6], Shang-Ran Huang[7], Yuan-Ren Cheng[7], Fu-Shun Hsu[7,8,9]

[1] Department of Biomedical Engineering, National Taiwan University, Taipei, Taiwan

[2] Department of Otolaryngology, Head and Neck Surgery, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Taipei, Taiwan

[3] Tzu Chi University, School of Medicine, Hualien, Taiwan

[4] Division of Sleep Surgery, Department of Otolaryngology, Head and Neck Surgery, Stanford University School of Medicine, Stanford, California, USA

[5] All Vista Healthcare Center and Center for AI and Advanced Robotics, National Taiwan University, Taipei, Taiwan

[6] Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan

[7]Heroic Faith Medical Science Co., Ltd., New Taipei, Taiwan

[8] Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

[9] Department of Critical Care Medicine, Far Eastern Memorial Hospital, New Taipei, Taiwan

Objective This study investigates the application of machine/deep learning to classify snoring sounds excited at single- or multi-levels of the upper airway in patients with obstructive sleep apnea (OSA) using data from drug-induced sleep endoscopy (DISE).
Methods The snoring sounds of 39 subjects were recorded during DISE procedures between November 2019 and November 2021. The snoring sounds were later analyzed and labeled according to the Velum (V), Oropharynx (O), Tongue Base (T), and Epiglottis (E) (VOTE) classification system based on DISE findings. The dataset, comprising 5,173 one-second segments, had V (n = 259), O (n = 403), T (n = 77), E (n = 13), VO (n = 1016), VT (n = 46), OT (n = 140), OE (n = 39), VOT (n = 30), and None (n = 3150) labels. We trained support vector machine (SVM), bidirectional long short-term memory (BiLSTM) networks, and ResNet-50 networks for obstruction site classification based on the snoring sounds with VOTE annotations.
Results The ResNet-50 showed the best overall performance in classifying snoring acoustics. The triplets of unweighted precision, recall, and F1-score for the SVM, BiLSTM, and ResNet-50 models were {0.86, 0.60, and 0.66}, {0.72, 0.67, and 0.69}, and {0.81, 0.71, and 0.75}, respectively.

Conclusion This study emphasizes the potential of integrating snoring acoustics with deep learning to localize multi-level upper airway obstruction. However, challenges such as limited sample size, data imbalance, and differences between pharmacologically induced and natural snoring sounds were noted, suggesting further research to enhance model accuracy and generalizability.

中文題目：採用深度學習分析打鼾聲音來定位上呼吸道多層次阻塞位置

作　　者：許瑛倢[1,2,3,*], 柳勇全[4], 黃昭蓉[4,5], 吳綺緯[6], 鄭人愷[6], 許永真[6], 黃上然[7], 鄭淵仁[7], 許富舜[7]

服務單位：

1 Department of Biomedical Engineering, National Taiwan University, Taipei, Taiwan

2 Department of Otolaryngology, Head and Neck Surgery, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, Taipei, Taiwan

3 Tzu Chi University, School of Medicine, Hualien, Taiwan

4 Division of Sleep Surgery, Department of Otolaryngology, Head and Neck Surgery, Stanford University School of Medicine, Stanford, California, USA

5 All Vista Healthcare Center and Center for AI and Advanced Robotics, National Taiwan University, Taipei, Taiwan

6 Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan

7 Heroic Faith Medical Science Co., Ltd., New Taipei, Taiwan

8 Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

9 Department of Critical Care Medicine, Far Eastern Memorial Hospital, New Taipei, Taiwan