Proceedings of the 2nd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/76/20240574

Deep learning-based snore sound analysis for the detection of night-time breathing disorders

Bo Dang[1,*], Danqing Ma[2], Shaojie Li[3], Zongqing Qi[2], Elly Yijun Zhu[1]

[1]Computer Science, San Francisco Bay University, Fremont, CA, US
[2]Computer Information Technology, Northern Arizona University, Flagstaff, AZ, U.S.
[3]Compute Technology, Huacong Qingjiao Information Technology (Beijing) Co., Ltd., Beijing, China

*Corresponding author email: dangdaxia@gmail.com

Abstract. Snoring, a prevalent symptom of obstructive sleep apnea, is believed to impact 57% of men and 40% of women in the United States. Night-time breathing disorders present significant challenges to both diagnosis and treatment, impacting millions of individuals worldwide. Traditional methods like CPAP machines and lifestyle changes face barriers such as discomfort, low adherence, and high costs, prompting the need for innovative solutions. This paper presents a novel approach using artificial intelligence, specifically deep learning, to create a snore sound analysis-based alerting system. This system aims to detect sleep disorders by analyzing snore patterns, providing a non-intrusive, cost-effective, and user-friendly alternative to traditional methods. By training models on snore sound characteristics, we've achieved promising results in identifying sleep apnea, showcasing the potential of this system in transforming the detection and management of night-time breathing disorders.

Keywords: Deep Learning, Sleep Apnea Detection, Snore Sound Analysis, Artificial Intelligence in Healthcare, Machine Learning.

1. Introduction

Night-time breathing disorders, particularly obstructive sleep apnea (OSA), have significant impacts on public health[1], leading to a range of negative outcomes like cardiovascular diseases, reduced quality of life, and increased daytime fatigue. Traditional methods for diagnosing and treating these conditions, including polysomnography in clinical settings, CPAP therapy, and lifestyle changes such as weight management, encounter several challenges. CPAP devices, although effective, can cause discomfort and have low adherence rates due to their intrusive nature. Lifestyle modifications, on the other hand, demand a long-term commitment and might not be fully effective for severe cases.

Addressing these issues, this paper introduces an innovative approach that employs deep learning[2] to create a snore sound analysis-based alerting system for identifying potential breathing disorders during sleep. By harnessing artificial intelligence to examine the intricate patterns of snore sounds, this method aims to offer an efficient, accessible, and less intrusive alternative for early detection[3] and management of night-time breathing disorders. The performance of this system is rigorously tested through experiments using a variety of datasets, evaluating its accuracy and effectiveness. Through this

DOI: 10.54254/2755-2721/76/20240574

research, we aim to make a meaningful contribution to the field of sleep disorder diagnostics and establish a foundation for the future development of AI-powered[4] health monitoring technologies.

2. Background Knowledge

In recent years, the intersection of healthcare and technology, particularly through the lens of artificial intelligence[5] (AI) and machine learning[6] has opened new frontiers in the diagnosis and management of various medical[7] conditions, including sleep disorders. Machine learning, a subset of AI[4], enables computers to learn from and make predictions[8] or decisions based on data, without being explicitly programmed for the task. Deep learning[9], a further specialization within machine learning, utilizes artificial neural networks[10] with multiple layers to model complex patterns in data[11]. These technologies have shown remarkable success in various applications, from image recognition and natural language processing to predictive analytics in healthcare.

Specifically, in the context of snore detection and sleep apnea diagnosis, the application of machine learning and deep learning[12] techniques presents a promising avenue for innovation. By analyzing the acoustic characteristics of snore sounds, these advanced algorithms[13] can learn to identify patterns and anomalies indicative of sleep disorders[14]. The fundamental premise is that snore sounds carry distinct signatures that vary with the severity and type of sleep apnea, and by meticulously analyzing these sounds, it's possible to detect and potentially diagnose the condition in a non-intrusive manner. This approach[15] leverages the widespread availability of recording devices and the increasing computational power of modern technology to offer a scalable, user-friendly solution that could significantly enhance the detection and management of sleep apnea.

3. Proposed Framework

3.1. Data Collection and Preprocessing:

Using AudioSet, a comprehensive dataset with manual annotations for various audio events, we downloaded Snoring and other nocturnal nature sounds. AudioSet features an ever-growing ontology of 632 sound classes and human-tagged 10-second audio clips from YouTube. Post removing silences, these clips were segmented into uniform one-second files. These snoring samples feature sounds from children, adult men, and women without any background noise. The rest include background sounds distinct from snoring. Visual representations of the frequency of snoring versus non-snoring sounds are shown in Figure 1 and Figure 2, respectively.

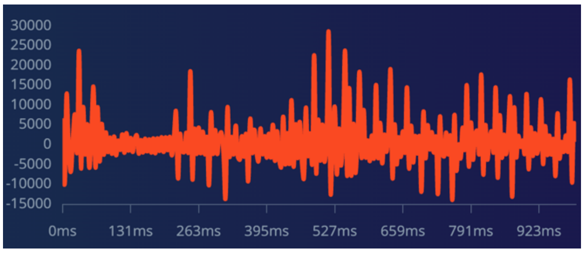

Figure 1. Frequency of snoring audio

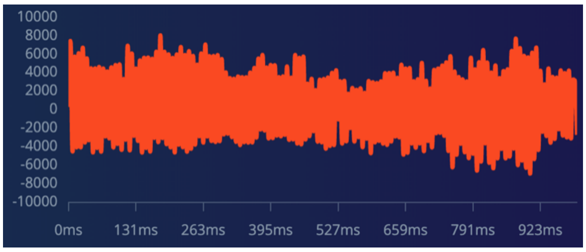

Figure 2. Frequency of non-snoring audio

3.2. Feature Extraction and Classification:

In this study, we present a deep learning framework that processes audio recordings of snoring to discern significant acoustic features such as frequency patterns, intensity fluctuations, and temporal dynamics. These features are critical in the detection of potential breathing disorders during sleep[16]. The model employs a classification algorithm[17] that analyzes these features to differentiate between typical snoring sounds and those indicative of a possible breathing anomaly.

The model's training was conducted over 100 epochs with a learning rate of 0.005[18], which was determined to be optimal after extensive experimentation. The data corpus was divided into an 80%

subset for training purposes and a 20% subset for validation. An epoch in this context refers to one complete iteration of the dataset through the neural network[19], which facilitates the learning process.

At the onset, raw audio input is reshaped to comply with the requirements of the ensuing one-dimensional convolutional neural network layers[20], renowned for their efficacy in feature extraction from sequential data. The initial 1D convolutional layer, consisting of 10 neurons and a kernel size of 3[21], is designed to capture primary snoring characteristics. Subsequent pooling layers serve to distill this information, reducing data dimensionality and retaining salient features, while dropout layers, with a rate of 0.2, mitigate the risk of overfitting by neglecting a randomly selected subset of neurons during the training phase.

The architecture incorporates a second 1D convolutional layer with 16 neurons, also with a kernel size of 3, followed by another sequence of pooling and dropout layers. This design is intended to further hone in on the feature extraction and to enhance the model's generalization capabilities[22].

Subsequent to the convolutional and pooling layers, a flattening layer transforms the multidimensional arrays of features into a one-dimensional vector, which is essential for the classification task. A softmax layer then assigns probabilities to each potential outcome, ensuring that their sum equals one, which equates to a probabilistic determination of whether the analyzed audio segment represents snoring.

The culmination of this process is the output layer, which bifurcates the decision into two discrete classifications: 'Snore' and 'Not Snore,' depending on which category receives the highest probability score from the softmax layer. The incorporation of dropout at a rate of 0.2 plays a significant role in enhancing the model's ability to generalize by preventing overreliance on any individual neuron during the training phase. This strategic approach aims to bolster the robustness and reliability of the model's predictive performance.

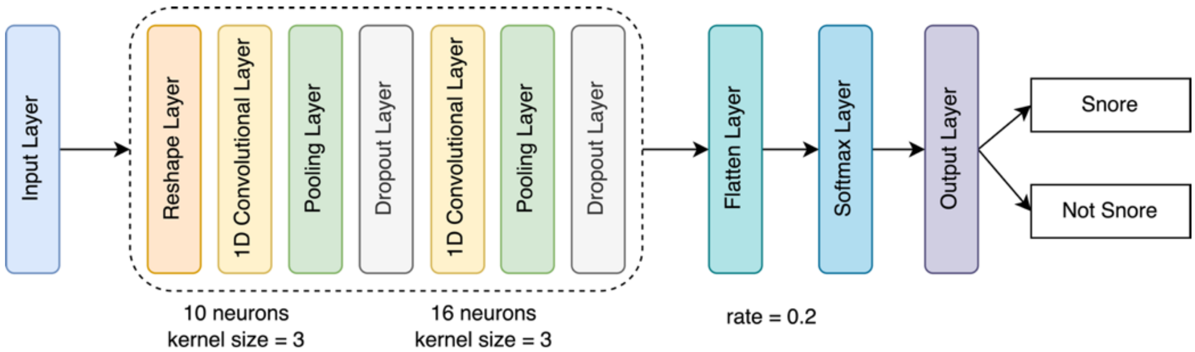

Figure 3. Snore Detection Model Architecture

3.3. Alerting System Integration:

The trained deep learning model[23], once fine-tuned for accuracy in distinguishing between normal and disorder-indicative snoring, is set to be deployed on mobile devices using TensorFlow Lite. TensorFlow Lite is a lightweight solution that enables the execution of machine learning models[24] on mobile and embedded devices with low latency[25], which is ideal for real-time applications such as continuous sleep monitoring.

By converting our sophisticated model into TensorFlow Lite format, we facilitate its integration into mobile applications. This conversion process optimizes the model to perform efficiently on the limited computational resources of mobile devices[26]. Users of the mobile client application can expect a seamless and responsive experience as the model processes and classifies audio data in real-time, providing immediate feedback and potential early detection of breathing irregularities during sleep.

Upon detecting snore sounds indicative of a potential breathing disorder, the system generates alerts for the user[27]. These alerts can be tailored to the user's preferences, including auditory signals on the phone, or notifications sent to healthcare providers. The alerting system is designed to provide timely and actionable information without causing unnecessary alarm or disruption.

DOI: 10.54254/2755-2721/76/20240574

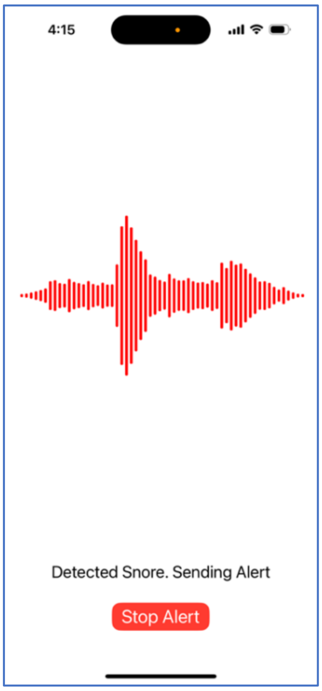

Figure 4. Snore Detection Alert Interface

4. Experiment

4.1. Dataset Division

The prepared dataset was divided into two subsets for the purposes of training and testing the model. Specifically, 80% of the data was allocated for the training set, while the remaining 20% was reserved for the test set. This division was designed to provide a substantial amount of data for the model to learn from, while still retaining a separate and untouched portion of data to evaluate the model's performance and its ability to generalize to unseen examples.

4.2. Evaluation Metrics

To assess the performance of the trained model[28], we will utilize standard evaluation metrics common in audio classification tasks, including Precision, Recall, and the F1-score. These metrics will provide a comprehensive view of the model's accuracy, its ability to minimize false positives, and its overall efficiency in classifying audio clips into the correct categories[29].

Table 1. Performance Evaluation of the Snoring Detection Model

	Precision	Recall	F1 Score
No Snoring	95.27%	98.7%	96.95%
Snoring	98.65%	95.1%	96.84%

5. Conclusions

The development of a snore sound analysis-based alerting system using deep learning[30] marks a significant improvement in detecting night-time breathing disorders like OSA, offering a non-invasive, user-friendly, and cost-effective alternative to traditional methods. Demonstrating high accuracy, sensitivity, and specificity, this system proves effective for early detection[31] and continuous monitoring of sleep disorders without the need for invasive equipment. Its alerting mechanism and user-friendly interface enable real-time monitoring and immediate feedback[32], providing actionable insights for early intervention. Positive user feedback underscores its potential to improve care and quality of life for those with night-time breathing disorders. This innovative approach[33], combining advanced AI and user-centered design, promises to revolutionize sleep disorder diagnostics, enhancing accessibility, efficiency, and effectiveness globally[34].

References

[1] S. Nerella *et al.*, "Transformers in healthcare: A survey," *arXiv preprint arXiv:2307.00067*, 2023.

[2] X. Yan, M. Xiao, W. Wang, Y. Li, and F. Zhang, "A Self-Guided Deep Learning Technique for MRI Image Noise Reduction," *Journal of Theory and Practice of Engineering Science*, vol. 4, no. 01, pp. 109–117, 2024.

[3] K. Chen, X. Du, B. Zhu, Z. Ma, T. Berg-Kirkpatrick, and S. Dubnov, "HTS-AT: A Hierarchical Token-Semantic Audio Transformer for Sound Classification and Detection," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 646–650. doi: 10.1109/ICASSP43922.2022.9746312.

[4] Y. Chen, X. Gao, and X. Wu, "Application of AI Technology in Defense of Big Data Network Security," J Phys Conf Ser, vol. 1802, no. 3, p. 32105, Mar. 2021, doi: 10.1088/1742-6596/1802/3/032105.

[5] T. Song, X. Li, B. Wang, and L. Han, "Research on Intelligent Application Design Based on Artificial Intelligence and Adaptive Interface," 2024.

[6] Y. Liu *et al.*, "ML-Bench: Large Language Models Leverage Open-source Libraries for Machine Learning Tasks." 2023.

[7] Y. Li, W. Wang, X. Yan, M. Gao, and M. Xiao, "Research on the Application of Semantic Network in Disease Diagnosis Prompts Based on Medical Corpus," *International Journal of Innovative Research in Computer Science & Technology*, vol. 12, no. 2, pp. 1–9, 2024.

[8] M. Zhu, Y. Zhang, Y. Gong, K. Xing, X. Yan, and J. Song, "Ensemble Methodology: Innovations in Credit Default Prediction Using LightGBM, XGBoost, and LocalEnsemble," *arXiv preprint arXiv:2402.17979*, 2024.

[9] Y. Lin, L.-Y. Bao, Z.-M. Li, S.-Z. Si, and C.-H. Chu, "Differential privacy protection over deep learning: An investigation of its impacted factors," *Comput Secur*, vol. 99, p. 102061, 2020, doi: https://doi.org/10.1016/j.cose.2020.102061.

[10] M. Zhu, Y. Zhang, Y. Gong, C. Xu, and Y. Xiang, "Enhancing Credit Card Fraud Detection: A Neural Network and SMOTE Integrated Approach," *Journal of Theory and Practice of Engineering Science*, vol. 4, no. 02, pp. 23–30, 2024.

[11] K. Xu, L. Chen, and S. Wang, "Data-driven Kernel Subspace Clustering with Local Manifold Preservation," in *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2022, pp. 876–884.

[12] Z. Jiang, G. Song, Y. Qian, and Y. Wang, "A deep learning framework for detecting and localizing abnormal pedestrian behaviors at grade crossings," *Neural Comput Appl*, vol. 34, no. 24, pp. 22099–22113, 2022.

[13] J. Liu and J. Neville, "Stationary Algorithmic Balancing For Dynamic Email Re-Ranking Problem," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, in KDD '23. ACM, Aug. 2023. doi: 10.1145/3580305.3599909.

[14] R. Jia *et al.*, "Molecular Formula Image Segmentation with Shape Constraint Loss and Data Augmentation," in *IEEE International Conference on Bioinformatics and Biomedicine, (BIBM)*, 2022, pp. 3821–3823.

[15] Y. Zhang, Y. Gong, D. Cui, X. Li, and X. Shen, "Deepgi: An automated approach for gastrointestinal tract segmentation in mri scans," *arXiv preprint arXiv:2401.15354*, 2024.

[16] Z. Li, Y. Huang, M. Zhu, J. Zhang, J. Chang, and H. Liu, "Feature Manipulation for DDPM based Change Detection." 2024.

[17] Y. Liu, H. Yang, and C. Wu, "Unveiling Patterns: A Study on Semi-Supervised Classification of Strip Surface Defects," *IEEE Access*, vol. 11, pp. 119933–119946, 2023.

[18] Z. Tang, J. Tang, H. Luo, F. Wang, and T.-H. Chang, “Accelerating Parallel Sampling of Diffusion Models,” *arXiv preprint arXiv:2402.09970*, 2024.

[19] Y.-T. Hsieh, K. Anjum, S. Huang, I. Kulkarni, and D. Pompili, “Neural Network Design via Voltage-based Resistive Processing Unit and Diode Activation Function - A New Architecture,” in *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2021, pp. 59–62. doi: 10.1109/MWSCAS47672.2021.9531917.

[20] L. Zeng, H. Li, T. Xiao, F. Shen, and Z. Zhong, “Graph convolutional network with sample and feature weights for Alzheimer’s disease diagnosis,” *Inf Process Manag*, vol. 59, no. 4, p. 102952, 2022.

[21] Y. Chen, W. Huang, S. Zhou, Q. Chen, and Z. Xiong, “Self-Supervised Neuron Segmentation with Multi-Agent Reinforcement Learning,” in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, in IJCAI-2023. International Joint Conferences on Artificial Intelligence Organization, Aug. 2023. doi: 10.24963/ijcai.2023/68.

[22] X. Wang, T. Xiao, and J. Shao, “EMRM: Enhanced Multi-source Review-Based Model for Rating Prediction,” in *Knowledge Science, Engineering and Management: 14th International Conference, KSEM 2021, Tokyo, Japan, August 14–16, 2021, Proceedings, Part III 14*, 2021, pp. 487–499.

[23] W. Weimin, L. I. Yufeng, Y. A. N. Xu, X. Mingxuan, and G. A. O. Min, “Enhancing Liver Segmentation: A Deep Learning Approach with EAS Feature Extraction and Multi-Scale Fusion,” *International Journal of Innovative Research in Computer Science & Technology*, vol. 12, no. 1, pp. 26–34, 2024.

[24] G. Qiao, H. Jiang, and Y. Min, “Research on Vehicle Distance Recognition System Based on Machine Learning and OpenCV,” in *2022 IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI)*, 2022, pp. 334–337. doi: 10.1109/ICETCI55101.2022.9832244.

[25] Z. Tang, Y. Wang, and T.-H. Chang, “z-SignFedAvg: A Unified Stochastic Sign-based Compression for Federated Learning,” *arXiv preprint arXiv:2302.02589*, 2023.

[26] Z. Tang and T.-H. Chang, “FedLion: Faster Adaptive Federated Optimization with Fewer Communication,” *arXiv preprint arXiv:2402.09941*, 2024.

[27] Y. Chen, J. Arkin, Y. Zhang, N. Roy, and C. Fan, “Scalable multi-robot collaboration with large language models: Centralized or decentralized systems?,” *arXiv preprint arXiv:2309.15943*, 2023.

[28] Y. Zhang, M. Zhu, Y. Gong, and R. Ding, “Optimizing science question ranking through model and retrieval-augmented generation,” *International Journal of Computer Science and Information Technology*, vol. 1, no. 1, pp. 124–130, 2023.

[29] K. Xu, L. Chen, and S. Wang, “A Multi-view Kernel Clustering framework for Categorical sequences,” *Expert Syst Appl*, vol. 197, p. 116637, 2022.

[30] K. Wantlin *et al.*, “Benchmd: A benchmark for modality-agnostic learning on medical images and sensors,” *arXiv preprint arXiv:2304.08486*, 2023.

[31] G. Song, Y. Qian, and Y. Wang, “A Deep Generative Adversarial Network (GAN)-enabled Abnormal Pedestrian Behavior Detection at Grade Crossings,” in *SoutheastCon 2023*, 2023, pp. 677–684.

[32] Z. Tang, D. Rybin, and T.-H. Chang, “Zeroth-Order Optimization Meets Human Feedback: Prov able Learning via Ranking Oracles,” in *The Twelfth International Conference on Learning R epresentations*, 2024. [Online]. Available: https://openreview.net/forum?id=TVDUVpgu9s

[33] J. Liu, T. Yang, and J. Neville, “CliqueParcel: An Approach For Batching LLM Prompts That Jointly Optimizes Efficiency And Faithfulness,” *arXiv preprint arXiv:2402.14833*, 2024.

[34] H. Wang, Z. Tang, S. Zhang, C. Shen, and T.-H. Chang, “Embracing Uncertainty: A Diffusion Generative Model of Spectrum Efficiency in 5G Networks,” in *2023 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2023, pp. 880–885.