SerenEEG: A Lightweight Hybrid Machine Learning Approach to Insomnia Detection from Ear-EEG.

Maggie Sun and Hannah Tay

I. INTRODUCTION

A. Issue Statement

Insomnia is a widespread public-health challenge affecting 30-40% of the global population, with an additional 10-15% experiencing it chronically [1]. Beyond sleep disruption, chronic insomnia is strongly associated with mental-health conditions; individuals with persistent insomnia face up to 17 times higher risk of developing mental health disorders such as anxiety and depression [2]. In addition to its psychological burden, insomnia significantly hinders daytime academic and workplace performance due to impaired attention and memory.

B. Current Solutions and Associate Issues

Despite insomnia's prevalence and impact, effective treatment remains inaccessible to many. First-line interventions such as cognitive behavioral therapy require structured sessions with trained therapists and sustained patient participation, limiting scalability and access [3]. Pharmacological treatments, including sedative-hypnotics, often take 30-60 minutes to take effect and are associated with side effects such as next-day drowsiness, confusion, dry mouth, and constipation [4], [5]. Commercial cranial electrotherapy stimulation such as Alpha-Stim AID and Fisher-Wallace Stimulator cost from $600 to $1300 and show inconsistent clinical efficacy. Notably, a large multicenter randomized controlled trial of Alpha-Stim AID found that while the device was safe, it was no more effective than a sham device [6].

C. Reasoning

Since insomnia and anxiety are strongly associated with neurophysiological changes in brain activity, EEG provides a non-invasive method for capturing these changes, and while traditional scalp EEG provides high-quality neural signals, it is impractical for long term sleep monitoring due to its complexity and discomfort [7]. Ear-EEG offers a more portable and user-friendly alternative; however, large, well-annotated ear-EEG datasets for insomnia remain limited. By contrast, scalp EEG datasets provide extensive, high-quality data that capture signatures of sleep disruption [8]. Integrated with artificial intelligence, ear-EEG grows to be a promising solution to chronic insomnia, even more powerful than traditional scalp-EEG.

II. RELATED WORKS

A. Summary of Literature Review

Insomnia detection research over the past five years follows several major directions. One line of work has explored ear-EEG as an alternative to traditional scalp-EEG, with multiple studies reporting its efficacy in long-term sleep recording despite reduced signal amplitude and increased noise [9], [8], [10]. A second direction involves the integration of machine learning for detection, broadly falling into two paths: classical machine learning and deep learning. While classical models have consistently achieved high reported accuracies of over 90%, deep learning performance has varied between 80-90% [11], [12], [13], [14]. Not only have deep learning models failed to surpass classical ML, they continue to face limitations associated with black-box predictions. This trend is likely influenced by the limited amount of diverse datasets and the inherent difficulty of collecting data. Among these studies, only Tiwari et al. 2022 collected original EEG data while subsequent work relied exclusively on public datasets. Furthermore, none of these models have been evaluated in real-time clinical settings, with most studies limiting claims to potential applicability rather than demonstrated deployment. A third, parallel line of research has focused on the use of taVNS in mediating insomnia. Recent studies report that taVNS is safe, well-tolerated, and produces sustained improvements in symptoms for up to twenty weeks; however, further investigation is still required to clarify underlying mechanisms [15], [16].

B. Research Gaps

Although some progress has been made in insomnia detection using EEG, several critical research gaps remain. First, the availability of large, well-annotated insomnia datasets is limited, which constrains the deployment and robustness of both classical and deep learning models, which fail to leverage their full ability to capture EEG patterns. Collecting such datasets is particularly challenging due to the high cost, time requirements, need for overnight PSD and logistical barriers associated with long-term sleep monitoring. Second, the majority of prior studies focus on offline analysis rather than real-time deployment, limiting their clinical applicability particularly for wearable systems where latency and computational efficiency are critical. Third, while some neural networks have been explored, there remains a wide range of architectures that are under-investigated due to dataset size constraints, and many existing approaches often fail to outperform classical counterparts and address their black-box nature.

C. Proposed Solution

To address these gaps, SerenEEG introduces a hybrid system that integrates Random Forest (RF) and a 1D Convolutional

Neural Network (CNN). The CNN operates directly on preprocessed ear-EEG segments to learn compact latent representations, which are subsequently used by the classical models to perform subject-level classification. EEG data acquired from the SerenEEG hardware is processed in real time by this hybrid pipeline, and the resulting predictions are used to determine whether a taVNS intervention should be triggered. To address concerns surrounding model interpretability, SHAP-based analysis is applied to the classical classifier, reducing the black-box nature of the overall system.

III. METHODOLOGY

A. Data

EEG data were obtained from publicly available sleep and mental-health datasets, including ISRUC-Sleep, Mendeley EEG/EOG/EMG data from a cross-sectional study on psychophysiological insomnia and normal sleep subjects, and the Cyclic Alternating Pattern (CAP) Sleep Database. For classical machine learning experiments, the combined dataset comprised 122 subjects, including 11 healthy controls and 111 individuals with insomnia-related conditions. EEG signals from channels F4A1, C4A1, and O2A1 were used, all recorded using standard clinical scalp configurations. For deep learning experiments, the data expanded to include an additional 9 disordered subjects from the CAP Sleep database, resulting in a total of 131 subjects. EEG recordings were segmented into 30-second epochs sampled at 200 Hz. Each sample was labeled as 0 (minority class) or 1 (majority class). The resulting input tensor had dimensions of

$$X \in \mathbb{R}^{N_{\text{samples}} \times N_{\text{channels}} \times T}$$

To bridge the gap between scalp and ear-EEG modalities, SerenEEG trains models on scalp EEG channels referenced to the A1 mastoid, which is anatomically proximal to the ear and commonly used as a reference in both clinical EEG and ear-EEG configurations. This strategy enables scalable training on large scalp EEG datasets while maintaining compatibility with real-time ear-EEG deployment. Additionally, to support deployment on ear-EEG hardware and reduce domain mismatch, the Ear-EEG Sleep Monitoring 2019 (EESM19) dataset was additionally used for domain adaptation. EESM19 consists of multi-night ear-EEG recordings from healthy subjects collected in home environments. A subset of ear-EEG sleep recordings was processed using bipolar channel configurations and identical preprocessing steps to fine-tune the CNN while preserving its learned temporal representations. Table I summarizes the datasets, participant counts, and their usage across modeling pipelines.

B. Data Preprocessing

Raw EEG signals were first denoised using a bandpass filter from 0.5-40 Hz to remove low-frequency drift and high-frequency noise. Signals were then resampled to 200 Hz, segmented into 30-second epochs, and normalized per epoch using z-score normalization. Table II summarizes the preprocessing pipeline for all experiments. This preprocessing pipeline follows standard practices in sleep EEG analysis.

TABLE I
DATASET SUMMARY USED FOR TRAINING AND EVALUATION.

Dataset	Participants	Epoch Size	Used In
ISRUC-Sleep	100	30 s @ 200 Hz	DL, ML
Mendeley (1)	11	30 s @ 200 Hz	DL, ML
Mendeley (0)	11	30 s @ 200 Hz	DL, ML
CAP Sleep	9	30 s @ 200 Hz	DL
EESM19	10	30 s @ 200 Hz	DL

TABLE II
EEG PREPROCESSING PIPELINE APPLIED ACROSS ALL EXPERIMENTS.

Stage	Description
Filtering	0.5–40 Hz Butterworth bandpass to remove drift and high-frequency noise
Resampling	All recordings resampled to 200 Hz
Epoching	Segmentation into 30-second epochs across three EEG channels
Normalization	Z-score normalization applied per epoch

Feature extraction was only applied for classical machine learning models, while deep learning models trained directly on normalized data. For classical machine learning experiments, features were extracted from the preprocessed. These included five power spectral density bands (alpha beta delta gamma theta), four band ratios (delta/theta, delta/alpha, alpha/beta, theta/alpha), Hjorth parameters (mobility, complexity, variance), and three basic statistical features including mean, standard, and skew.

C. ML Development

1) Base-level Classical Machine Learning Models: As an initial modeling baseline, RF and XGB classifiers were trained directly on fifteen hand-crafted EEG features extracted from scalp EEG recordings.To improve generalization and reduce feature redundancy, Recursive Feature Elimination (RFE) was applied as part of the optimization pipeline. Hyperparameters were tuned using a two-stage strategy: RandomizedSearchCV was first used to identify promising regions of the search space, followed by GridSearchCV for fine-grained optimization. All evaluations were conducted using subject-wise cross-validation to prevent data leakage across epochs. These base-level models serve two purposes: first, they provide a transparent and well-established performance reference grounded in prior insomnia detection literature; second, they enable direct comparison between traditional feature-based approaches and the proposed deep learning–driven pipeline. Maintaining these base models allows clear attribution of performance gains to representation learning and domain adaptation rather than dataset or evaluation changes. This is particularly important given the class imbalance and limited number of subjects commonly encountered in sleep EEG studies.

2) CNN Model Architecture: A lightweight one dimensional convolutional neural network was developed to operate directly on preprocessed EEG epochs. The network consists of three convolutional blocks each with batch normalization
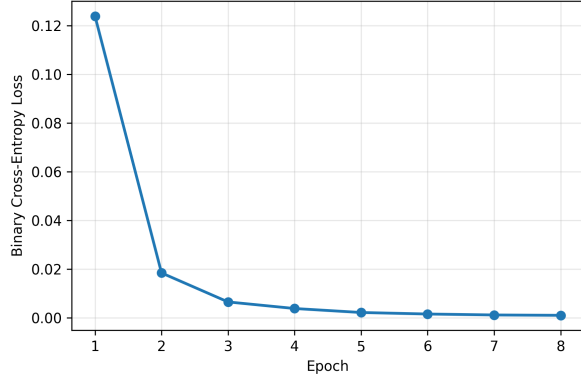
Fig. 1. Binary cross-entropy loss across fine-tuning epochs for domain adaptation of the CNN to EESM19. The rapid initial decrease followed by convergence suggests effective alignment of learned representations.

and max pooling, followed by global average pooling and fully connected layers. The CNN was trained using binary cross-entropy loss and optimized with the Adam optimizer. Dropout regularization was applied in the fully connected layers to mitigate overfitting, and the model outputs a single logit corresponding to the probability of insomnia-related sleep disturbance.

3) Domain Adaptation Using Ear-EEG Data: Since scalp EEG datasets and ear-EEG deployment have a domain gap, the CNN was first trained on scalp EEG data from ISRUC Sleep, Mendeley, and CAP Sleep Database and subsequently fine-tuned using EESM19. During fine-tuning, convolutional layers were frozen and only the fully connected layers were updated, allowing the model to adapt to ear-EEG signal characteristics while preserving learned temporal representations. Since the goal of domain adaptation is to align representations, loss was optimized and used instead of standard evaluation metrics due to its ability to answer the question of generalization and alignment. Figure III-C3 shows the decrease in loss over the course of domain adaptation.

4) CNN-Based embedding extraction and hybrid classification: Following fine-tuning, the CNN was repurposed as a feature extractor by removing the final classification layer. The resulting 64-dimensional embeddings capture high-level temporal representations learned from the raw EEG signals which were aggregated at the subject level using mean and standard deviation pooling to be used as inputs to RF and XGB classifiers. Final predictions are produced by RF classifier, while the CNN serves as a feature-learning backbone. Table ?? summarizes all experiments and their inputs representations.

TABLE III
MODELING STRATEGIES AND INPUT REPRESENTATIONS.

Model	Subjects	Input Representation	Samples
RF, XGB	122	Feature vectors (45D)	107,415
CNN	131	(3×6000) waveform	120,862
CNN + RF	131	CNN embeddings (64D)	120,862

5) Interpretability: To address interpretability concerns, SHAP (SHapley Additive exPlanations) analysis was applied to the RF and XGB models trained on CNN-derived embeddings. This enables identification of embedding dimensions most influential in model decisions, reducing the black-box nature of the system and providing post-hoc validation of learned representations.

REFERENCES

[1] J. Fernandez-Mendoza, "Insomnia Phenotypes, Cardiovascular Risk and Their Link to Brain Health," *Circulation Research*, vol. 137, no. 5, pp. 727–745, Aug. 2025.

[2] "How sleep affects mental health (and vice versa): What the science says," https://med.stanford.edu/news/insights/2025/08/sleep-mental-health-connection-what-science-says.html.

[3] "When You Can't Sleep: How CBT-I Can Help," https://www.massgeneral.org/news/article/cbt-i-for-sleep.

[4] "How Long Does Melatonin Take to Work?" Oct. 2022.

[5] "Prescription sleeping pills: What's right for you? - Mayo Clinic," https://www.mayoclinic.org/diseases-conditions/insomnia/in-depth/sleeping-pills/art-20043959.

[6] R. Morriss, S. Patel, C. Boutry, P. Patel, B. Guo, P. M. Briley, D. Butler, M. Craven, A. Duncan, C. Griffiths, F. Higton, R. McNaughton, N. Nixon, V. Prasad, K. Sayal, D. Smart, A. Zafar, and J. Kai, "Clinical effectiveness of active Alpha-Stim AID versus sham Alpha-Stim AID in major depression in primary care in England (Alpha-Stim-D): A multicentre, parallel group, double-blind, randomised controlled trial," *The Lancet Psychiatry*, vol. 10, no. 3, pp. 172–183, Mar. 2023.

[7] S. Hanzal, L. Tvrda, and M. Harvey, "An Investigation into Discomfort and Fatigue Related to the Wearing of an EEG Neurofeedback Headset," p. 2023.02.16.23284115, Feb. 2023.

[8] G. Hammour, H. Davies, G. Atzori, C. D. Monica, K. Ravindran, V. Revell, D.-J. Dijk, and D. Mandic, "From Scalp to Ear-EEG: A Generalizable Transfer Learning Model for Automatic Sleep Scoring in Older People," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 448–456, Apr. 2024.

[9] H. Moumane, J. Pazuelo, M. Nassar, J. Y. Juez, M. Valderrama, and M. Le Van Quyen, "Signal quality evaluation of an in-ear EEG device in comparison to a conventional cap system," *Frontiers in Neuroscience*, vol. 18, p. 1441897, Sep. 2024.

[10] A. S. Mihai (Ungureanu), O. Geman, R. Toderean, L. Miron, and S. SharghiLavan, "The Next Frontier in Brain Monitoring: A Comprehensive Look at In-Ear EEG Electrodes and Their Applications," *Sensors (Basel, Switzerland)*, vol. 25, no. 11, p. 3321, May 2025.

[11] S. Tiwari, D. Arora, and V. Nagar, "Detection of insomnia using advanced complexity and entropy features of sleep stage data of EEG recordings," *Measurement: Sensors*, vol. 24, p. 100498, Dec. 2022.

[12] M. Sharma, D. Anand, S. Verma, and U. R. Acharya, "Automated insomnia detection using wavelet scattering network technique with single-channel EEG signals," *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106903, Nov. 2023.

[13] S. Philip Mulamoottil and T. Vigneswaran, "A double-layered fully automated insomnia identification model employing synthetic data generation using MCSA and CTGAN with single-channel EEG signals," *Scientific Reports*, vol. 14, no. 1, p. 23427, Oct. 2024.

[14] M. Mondal, J. Papon, and M. Ahmad, "Automatic Classification of Insomnia Using Machine Learning Algorithms From EEG Signals," Jul. 2025, pp. 1–6.

[15] S. Zhang, Y. Zhao, Z. Qin, Y. Han, J. He, B. Zhao, L. Wang, Y. Duan, J. Huo, T. Wang, Y. Wang, and P. Rong, "Transcutaneous Auricular Vagus Nerve Stimulation for Chronic Insomnia Disorder: A Randomized Clinical Trial," *JAMA network open*, vol. 7, no. 12, p. e2451217, Dec. 2024.

[16] C. Liu, S. Chen, Y. Zhang, X. Wu, and J. Liu, "Transcutaneous Auricular Vagus Nerve Stimulation (taVNS) for Insomnia Disorder: A Narrative Review of Effectiveness, Mechanisms and Recommendations for Clinical Practice," *Nature and Science of Sleep*, vol. 17, pp. 1327–1344, Jun. 2025.