Real-Time and Energy-Efficient Attention Monitoring Using Single-Channel EEG for Everyday Life

ANICE JAHANJOO[1], VIMALA BAUER[2], SOHEIL KHOOYOOZ[2], MOSTAFA HAGHI[2], (Member, IEEE), NIMA TAHERINEJAD[2], (Member, IEEE)

[1]Institute of Computer Technology, TU Wien, Vienna, Austria(e-mail: firstname.lastname@tuwien.ac.at)
[2]Institute for Computer Engineering, Heidelberg University, Heidelberg, Germany (e-mail: firstname.lastname@ziti.uni-heidelberg.de)

Corresponding author: Anice Jahanjoo (e-mail: anice.jahanjoo@tuwien.ac.at).

ABSTRACT Attention monitoring is essential in tasks requiring sustained cognitive focus. Electroencephalography (EEG) offers a portable, cost-effective, and non-invasive solution; however, its signal complexity requires advanced processing. This study aims to develop an accurate and efficient method for classifying attention using a minimal EEG setup suitable for wearable applications. We propose an optimized single-channel EEG approach that combines Recursive Feature Elimination (RFE) for feature selection with a customized eXtreme Gradient Boosting (cXGB) classifier. The use of a single EEG electrode enables compatibility with commercially available, user-friendly EEG headbands, making real-time attention monitoring feasible in daily life. Our findings show that the proposed cXGB model achieves 98.29% binary and 94.25% three-class classification accuracy under subject-wise cross-validation, outperforming previous multi-channel approaches by 3%. Furthermore, the optimized feature set reduces processing time, memory usage, and energy consumption on both high-end (Intel i9) and low-power (Raspberry Pi) devices. This work provides an efficient and scalable solution for real-time, individualized attention monitoring on wearable platforms.

INDEX TERMS Attention detection, EEG signal processing, edge computing, feature engineering, machine learning, wearable devices.

I. INTRODUCTION

RAPID advances in wearable sensing have enabled portable measurement of mental and physiological states [1]. Electroencephalography (EEG) captures brain activity via scalp electrodes and is widely used. However, multi-channel systems are often obtrusive, complex, and impractical for daily use, limiting their adoption in real-world wearable applications. This underscores the need for accurate, lightweight, and user-friendly EEG-based approaches that can operate in real-time on wearable devices.

EEG is widely applied in clinical diagnostics (e.g., epilepsy [2], Attention Deficit and Hyperactivity Disorder (ADHD) [3]), brain-computer interfaces such as motor imagery [4], and cognitive or affective state monitoring [5]–[8]. In safety-critical and industrial settings, fluctuations in attention due to distraction or fatigue can degrade performance and increase risk [9], [10]. EEG provides a direct window into these fluctuations, offering potential for productivity and safety monitoring in demanding environments.

EEG rhythms are classically grouped into delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz), and beta (13-30Hz), each linked to cognitive states [11]. Multi-channel recordings combined with machine-learning classifiers have been effective for attention/distraction detection [12], [13], but require complex setups. Multi-channel headbands, although widely used in research [14], [15], remain impractical for real-world wearables. A complementary strategy is to focus on single electrodes over scalp regions where attention-related spectral changes are pronounced [12]. This reduces hardware burden while retaining discriminative signal content. The challenge is to design models that are accurate, data-efficient, and computationally frugal for deployment.

Recent advances in deep learning have introduced powerful models for EEG decoding, including Convolutional Neural

Network (CNN), recurrent architectures enhanced with attention mechanisms, Graph Neural Networks, and transformer-based architectures [16], [17]. While these approaches have achieved strong performance on large-scale datasets, they typically require substantial computational resources, large memory footprints, and high energy consumption, which restrict their suitability for real-time, edge, or wearable applications. In addition, such models often demand extensive hyperparameter tuning and large volumes of training data to generalize effectively, conditions that are not always feasible in attention-monitoring scenarios. In this work, we focus on accurate yet computationally efficient solutions; therefore, rather than employing complex deep architectures that can be more demanding than CNNs, we concentrate on boosting-based methods that provide a favorable balance between accuracy and resource efficiency. This approach allows our algorithm to be integrated with commercially available, user-friendly EEG headbands, which typically offer only a single electrode, enabling real-time attention monitoring in daily-life settings.

Gradient Boosting Decision Trees (GBDT) such as eXtreme Gradient Boosting (XGB) [18], [19] are attractive for EEG because of small memory footprint, fast inference, and interpretability [20], [21]. A limitation is that standard XGBoost uses fixed or monotonic learning-rate schedules, which may lead to underfitting or overfitting small, nonstationary EEG data. Motivated by adaptive schedules in deep learning [22], we investigate a hybrid adaptive learning-rate mechanism for boosting.

Using a public attention dataset [13], we present a single-channel pipeline that combines standard spectral features with a customized XGBoost configuration. In our experiments, the proposed customized eXtreme Gradient Boosting (cXGB) achieves the best accuracy among the evaluated classical baselines, while remaining competitive in prediction speed and energy consumption on edge hardware. We emphasize this study as a feasibility analysis for wearable use; broader generalization is limited by dataset size and is discussed in Section V.

The key contributions of this work are:

i) We propose a lightweight, single-electrode EEG pipeline for attention monitoring, specifically designed for wearable and edge devices, enabling real-time operation on user-friendly headbands with minimal hardware.

ii) We design a customized XGBoost model with a hybrid adaptive learning-rate schedule and class-weighted loss, which improves performance on EEG datasets while remaining computationally efficient for daily-life wearable applications.

iii) We provide a workflow from preprocessing and feature extraction to modeling and deployment, validated on a public dataset with runtime and energy measurements on both high-end and low-power platforms, demonstrating practical feasibility and scalability.

The paper is structured as follows: Section II provides a review of related work, and Section III outlines the methodology and key steps of our approach. Section IV presents the experimental results and comparative analysis, followed by Section V, which offers an in-depth discussion of the findings. Finally, Section VI concludes the paper.

II. RELATED WORK

Cognitive monitoring using EEG sensors and Machine Learning (ML) algorithms has been widely studied in contexts such as online learning [23], learning disabilities [24], and cognitive workload monitoring [25]. Jin et al. [26] employed Support Vector Machine (SVM) to predict mind-wandering with an accuracy of 60% using EEG data collected from 18 subjects performing the Sustained Attention to Response Task (SART) and a visual search task. Mohamed et al. [27] achieved an accuracy of 84% in detecting attention using 14-channel EEG signals collected from 86 participants. Al-Nafjan et al. [23] achieved an accuracy of 96% in predicting students' attention levels in online learning using Random Forest (RF) classifier. The authors employed frequency-domain features, including Power Spectral Density (PSD) of different frequency bands extracted from each of the seven channels of the EEG data. They applied ML classifiers such as K-Nearest Neighbors (KNN), SVM and RF, to recognize focused and unfocused mental states.

Suhail et al. [28] collected data using a 128-channel EEG device from 33 subjects to assess cognitive states. The authors extracted and analyzed 83 EEG features, including band ratios, relative power, and statistical features, to train classifiers. The model achieved a peak accuracy of 97.85% for binary classification between resting and focused states. Acı et al. [13] achieved an accuracy of 91% in detecting focused, unfocused, and drowsy states using SVM. The authors also made their EEG attention monitoring dataset publicly available on Kaggle, enabling further research in this area. Devi et al. [29] extracted features using a Long-Short-Term-Memory (LSTM) model and subsequently employed a graph-based conventional attention neural network (GA-CNN) model to classify students' cognitive states. Their proposed approach demonstrated improvements in binary classification, achieving an accuracy of 87%, which outperformed other existing methods.

Our study aims to enhance attention state detection accuracy while minimizing the number of EEG channels, unlike previous research that relied on multi-electrode EEG devices, which are impractical for daily use. By refining a ML algorithm and optimizing the number of channels without compromising accuracy, our one-channel method demonstrates practical feasibility for integrating into simple wearable devices, such as headbands, which are typically equipped with only one or a few strategically deployed electrodes. Furthermore, we evaluate the processing time, energy consumption, and memory efficiency of our approach on two hardware platforms to demonstrate its practicality for real-time edge device applications, an aspect frequently overlooked in previous studies.
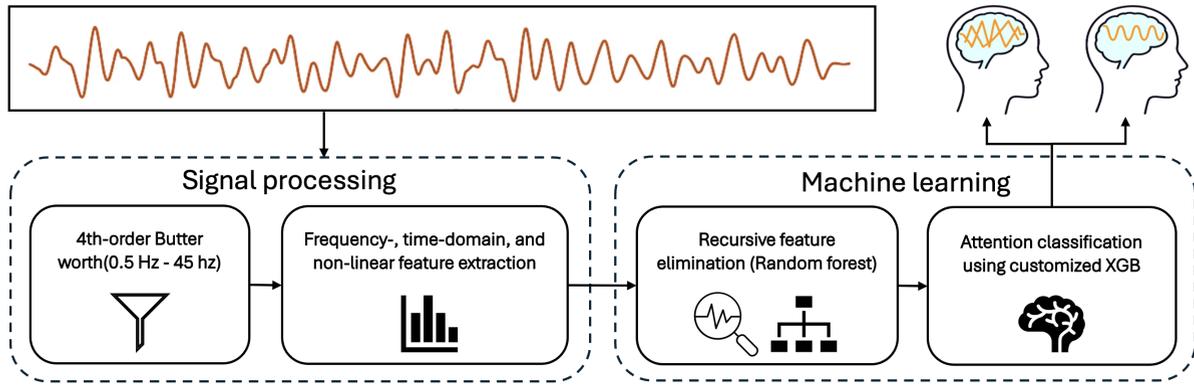
IEEE *Access*

FIGURE 1. Proposed pipeline for single-channel EEG attention classification

III. PROPOSED METHOD

A. EXPERIMENTAL SCENARIO

This study proposes a single-channel EEG-based attention detection framework designed for wearable applications. The overall system scenario is illustrated in Figure 1, which outlines the major stages of the process: signal preprocessing, feature extraction and selection, and ML-based classification. In the proposed setup, participants wear an EEG headset (Emotiv Epoc X) while engaging in a continuous monitoring task that simulates real-world attention demands in which participants supervised a virtual train using the Microsoft Train Simulator. This task requires sustained attention with minimal physical activity, resembling real-life conditions such as vehicle monitoring or control-room supervision. EEG signals are recorded from 14 electrodes in the original dataset, but for this study, we selected the AF3 electrode, located in the frontal region of the scalp, due to its sensitivity to attentional modulation and its relevance for single-channel, wearable applications. The recorded EEG data are first filtered to remove noise and artifacts using a 4th-order Butterworth bandpass filter. The cleaned signals are then segmented into fixed-length windows of non-overlapping 15 seconds, each representing a short-duration cognitive state suitable for near real-time analysis. From each segment, a set of statistical, spectral, and nonlinear features is extracted to capture the underlying neural dynamics associated with different attention levels. To improve model interpretability and efficiency, a Recursive Feature Elimination (RFE) process identifies the most informative features. The selected features are then used to train the proposed cXGB model, which incorporates a hybrid adaptive learning rate mechanism and weighted loss adjustment to classify the attention. This scenario reflects a practical and lightweight architecture for wearable EEG-based attention monitoring, emphasizing low-channel complexity, adaptive learning, and real-time feasibility, while remaining grounded in an experimentally validated setup.

B. DATASET

Publicly available EEG datasets with well-annotated attention labels are scarce, particularly for single-channel or reduced-channel configurations. Most large-scale EEG databases are designed for clinical diagnostics (e.g., epilepsy, seizure detection) or motor imagery, and they rarely provide explicit annotations of attention or cognitive state suitable for benchmarking attention classification models. Collecting a new large-scale dataset was beyond the scope of this work due to the substantial cost, time, and ethical requirements associated with EEG experiments. Therefore, we selected the benchmark dataset provided by Aci et al. [13], which, despite its relatively small subject pool, is one of the few publicly available datasets that directly targets attention-related cognitive states, provides controlled experimental conditions, and has been used in prior EEG-based attention studies. By relying on a publicly available resource, we also ensure the reproducibility and comparability of our findings with future research.

We used the dataset from Aci et al. [13], which contains 25 hours of EEG recordings from 5 participants performing a low-intensity control task using the Microsoft Train Simulator. Each experiment involved participants supervising a simulated passenger train along a flat, featureless route for 35-55 minutes. Participants simulated three mental states: focused attention, unfocused but awake, and drowsy, each lasting 10 minutes per session. In the focused state, participants concentrated on the simulation without active intervention. During the unfocused state, they disengaged from the task while remaining awake, representing a state challenging to detect through external cues. Finally, during the drowsy state, participants were allowed to relax and doze off freely. The train simulation settings required participants to maintain a steady speed of 64.4 km/h using basic throttle and brake controls on a keyboard. Each participant performed up to seven experiments, where the first two were intended for habituation and the remaining five for data collection. The experiments were conducted between 7 PM and 9 PM to facilitate drowsiness during the final phase. Participants were monitored via video recordings to ensure compliance with

the experimental protocol. The dataset and device details are summarized in Table 1.

For this dataset, EEG data was recorded using the Emotiv Epoc X device. Figure 2 demonstrates all EEG channels in the dataset represented by colored points (left) alongside the Emotiv Epoc X device (right). For this study, we selected the data from the AF3 channel, highlighted in pink, due to its placement over a region known for significant neural activity changes associated with attention [12]. EEG studies have demonstrated that changes in alpha (8-12 Hz) and theta (4-8 Hz) rhythms at frontal regions, including AF3, correlate with attentional states. Specifically, alpha suppression and theta increase are linked to heightened attention and focused engagement [30]. Furthermore, the AF3 channel is placed near Fp1, which is commonly used in most commercial EEG devices [31]–[33], ensuring compatibility and practicality for future applications.

C. SIGNAL PREPROCESSING AND SEGMENTATION
EEG signals are often contaminated with various artifacts, including baseline drift, muscle activity, eye movements, and electrical noise. To address these issues, a 4th-order Butterworth bandpass filter with a frequency range of 0.5 to 45 Hz were applied. The lower cutoff at 0.5 Hz removes slow drift artifacts typically caused by sweat or minor movements [34], while the upper cutoff at 45 Hz suppresses high-frequency noise from muscle activity and environmental electrical interference. This filtering ensures that the relevant EEG frequency bands are preserved for cognitive analysis.

Following preprocessing, the EEG signals were segmented into non-overlapping 15-second windows. This duration was chosen to balance analytical accuracy and practical feasibility. Each segment is long enough to capture multiple cycles of lower-frequency components such as theta and alpha bands, which are linked to attention, while remaining short enough to allow near-real-time analysis for wearable applications. Furthermore, segmenting the EEG data into 15-second windows is consistent with methodologies employed in related studies, facilitating direct comparison of our results with existing literature and established benchmarks.

D. FEATURE EXTRACTION AND SELECTION
A comprehensive set of 30 features was initially extracted from each EEG segment to capture multiple dimensions of

TABLE 2. List of Preliminary Features

Features (T: Time domain, F: Frequency domain, R: Relative, N: Nonlinear)	
1. PSD Mean (F)	16. Simple Square Integral (T)
2. Relative Alpha Power (F)(R)	17. Absolute Energy (F)
3. Relative Beta Power (F)(R)	18. Log Energy (F)
4. Entropy (T)	19. Differential Variance (T)
5. Mean (T)	20. Normalized Second Difference (T)(R)
6. Standard Deviation (T)	21. Normalized First Difference (T)(R)
7. Skewness (T)	22. Sum of Square Roots (T)
8. Kurtosis (T)	23. Exponential Root Sum (T)
9. Median Frequency (F)	24. Differential Absolute Standard Deviation (T)
10. Total Power (F)	25. Higuchi Fractal Dimension (N)
11. Trimean (T)	26. Hjorth Mobility (N)
12. Interquartile Range (T)	27. Average Energy (F)
13. Zero-Crossing Rate (T)	28. Wavelength (T)
14. Minima (T)	29. Hurst Exponent (N)
15. Peak Amplitude (T)	30. Slope Sign Change (T)

neural activity, including time-domain, frequency-domain, nonlinear, and relative measures. These features, summarized in Table 2, provide a representation of EEG signal characteristics relevant to attentional states.

To optimize model performance and reduce computational complexity, RFE was employed to identify the most informative subset of features. RFE iteratively removes the least significant features based on their contribution to model accuracy, continuing the process until the optimal subset is determined. We used a RF algorithm with 100 decision trees as the base estimator, and the step size was set to 1, eliminating one feature per iteration. This elimination process was repeated 85 times, progressively reducing the feature set from 30 to 20 features. Features that were consistently selected across iterations were retained, ensuring that the final subset robustly captured the EEG patterns most predictive of attention levels. This systematic selection not only improves classification performance but also enhances interpretability and reduces the risk of overfitting.

E. MACHINE LEARNING CLASSIFICATION
We perform both binary and three-class classification to detect attention levels from EEG data. In binary classification, we label the focused state as the attention class while grouping the unfocused and drowsy states into the non-attention class. To address the class imbalance, we utilize the scikit-learn [35] library to upsample the minority class, balancing the dataset. No resampling was applied for the three-

TABLE 1. EEG Dataset Description

Attribute	Description
Number of Classes	Three classes: focused, unfocused, and drowsiness.
EEG Device	Emotiv Epoc X (Emotiv, LA, California, USA) with a 128 Hz sampling frequency.
Number of Subjects	Five subjects, 33 sessions (seven sessions per subject except subject 5).
Duration	25 hours of EEG recordings, with sessions lasting approximately 30 minutes.

(a) Placement of EEG channels on the head.

(b) Emotiv Epoc X
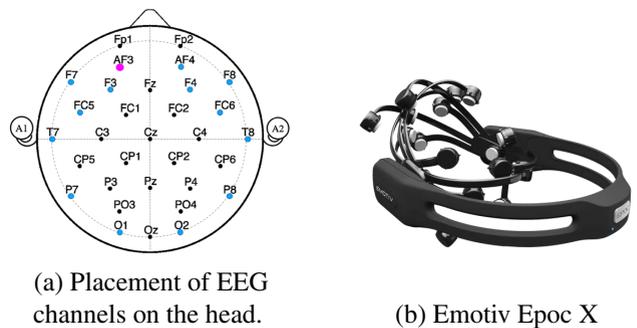
FIGURE 2. EEG channels in the dataset, with the selected AF3 channel highlighted in pink.

IEEE *Access*

class classification, which includes focused, unfocused, and drowsy states.

Our cXGB method addresses two critical challenges in EEG-based attention detection: (1) the dynamic nature of neural signals requiring adaptive learning strategies, and (2) the challenge of improving the generalizability of attention detection. The proposed solution combines novel learning rate adaptation with specialized loss weighting, as formalized in Algorithm 1.

Traditional GBDT methods employ static learning rates:

$$\hat{y}^{(t)} = \hat{y}^{(t-1)} + \eta f_t(x), \qquad (1)$$

where η remains fixed throughout training. In EEG analyses, where signal-to-noise ratios (SNR) show significant variability across trials and subjects, enhancing the generalizability of predictive models becomes crucial. To address this challenge, we propose integrating three dynamic learning rate components through a Hybrid Adaptive Learning Rate (HALR) mechanism. This approach aims to adaptively adjust the learning rate during training, thereby improving model performance and robustness across diverse EEG data.

Oscillating Learning Rate (OLR) periodically resets learning rates to escape local minima:

$$\eta_{\text{OLR}}^{(t)} = \eta_{\min} + 0.5(\eta_{\max} - \eta_{\min})\left(1 + \cos\left(\frac{\pi t}{T}\right)\right), \quad (2)$$

with $\eta_{\min} = 0.01$, $\eta_{\max} = 0.5$, and cycle length $T = 50$, inspired by cosine annealing cycles [36].

Gradient-Based Learning Rate (GBLR) adapts to signal volatility through:

$$\eta_{\text{GBLR}}^{(t)} = \max\left(\frac{\eta_{base}}{1 + \bar{g}^{(t)}}, 0.01\right), \qquad (3)$$

where $\eta_{\text{base}} = 0.1$ initializes the adaptation mechanism, and $\bar{g}^{(t)} = \frac{1}{N}\sum_{i=1}^{N}|\sigma(f_t(x_i)) - y_i|$ represents mean absolute gradients of the logistic loss.

Performance-Driven Adaptation monitors validation loss windows:

$$\eta^{(t)} = \begin{cases} 0.5\eta^{(t-1)} & \text{if } \Delta\mathcal{L}_{\text{val}}^{(t:t-5)} < 10^{-4} \\ \eta^{(t-1)} & \text{otherwise} \end{cases} \qquad (4)$$

These components combine through:

$$\eta^{(t)} = \text{clip}\left(\frac{\eta_{\text{GBLR}}^{(t)} + \eta_{\text{OLR}}^{(t)}}{2}, 0.01, 0.5\right), \qquad (5)$$

with clipping preventing instability during high-variance EEG epochs.

In addition, in our cXGB, we customize the XGB objective function to enhance precision in binary classification. Specifically, we introduce weighted penalties for False Positive (FP) and False Negative (FN), addressing residual class imbalance even after dataset balancing. This modification ensures a more equitable reduction in misclassification errors based on their practical importance.

$$\mathcal{L} = -w_{FN}y\log(\hat{y}) - w_{FP}(1-y)\log(1-\hat{y}), \qquad (6)$$

Algorithm 1 Pseudocode for HALR-based cXGB Training for EEG Attention Detection

---

1: **Input:** Training data $\{(x_i, y_i)\}_{i=1}^{N}$, max iterations T, base learning rate η_{base}, FP/FN weights w_{FP}, w_{FN}
2: **Output:** Predicted model $\hat{y}$
3: Initialize prediction: $\hat{y}^{(0)} \leftarrow 0$, learning rate: $\eta^{(0)} \leftarrow \eta_{\text{base}}$
4: **for** $t \leftarrow 1$ **to** T **do**
5:     Compute weighted gradient for each sample:
$$g_i = (\hat{y}_i^{(t-1)} - y_i) \cdot (y_i w_{FN} + (1 - y_i)w_{FP})$$
6:     Compute weighted Hessian for each sample:
$$h_i = \hat{y}_i^{(t-1)}(1 - \hat{y}_i^{(t-1)}) \cdot (y_i w_{FN} + (1 - y_i)w_{FP})$$
7:     Compute adaptive learning rate: $\eta_{\text{GBLR}}^{(t)}$ (Eq. 3)
8:     Compute oscillating learning rate: $\eta_{\text{OLR}}^{(t)}$ (Eq. 2)
9:     Combine into hybrid learning rate: $\eta^{(t)}$ (Eq. 5)
10:     Fit tree f_t to minimize $\sum_i[g_i f_t(x_i) + 0.5 h_i f_t^2(x_i)]$
11:     Update predictions: $\hat{y}^{(t)} \leftarrow \hat{y}^{(t-1)} + \eta^{(t)}f_t(x)$
12:     Evaluate validation loss $\mathcal{L}_{\text{val}}^{(t)}$
13:     **if** $\Delta\mathcal{L}_{\text{val}}^{(t:t-5)} < 10^{-4}$ **then**   ▷ PDLR (Eq. 4)
14:         $\eta^{(t)} \leftarrow 0.5 \cdot \eta^{(t)}$
15:     **end if**
16: **end for**
17: **return** $\hat{y}^{(T)}$

---

where $y \in \{0, 1\}$ is the true label, $\hat{y} = \sigma(f_t(x))$ represents the sigmoid-transformed predicted probability ($\sigma(z) = 1/(1+e^{-z})$), and w_{FN}, w_{FP} are weights for FN and FP respectively. To prioritize reducing FP and FN, we introduce custom *weights* in the *gradient* and *Hessian* of the loss function. We set the weights w_{FP} and w_{FN} to 0.66 and 0.34 respectively, reflecting the relative importance of minimizing FN and FP based on our dataset results. These weights are normalized so that their sum equals 1, ensuring a balanced contribution from both types of misclassification in the loss function. For the gradient, the standard formula is:

$$\text{grad}_i = \hat{y}_i - y_i. \qquad (7)$$

To incorporate the asymmetric importance of FNs and FPs, we adjust the gradient using the class-dependent weights:

$$\text{grad}_i = (\hat{y}_i - y_i) \times (y_i \cdot w_{FN} + (1 - y_i) \cdot w_{FP}). \qquad (8)$$

This ensures that errors in the positive class ($y_i = 1$) are scaled by w_{FN}, and errors in the negative class ($y_i = 0$) are scaled by w_{FP}. As a result, the boosting algorithm penalizes misclassifications in proportion to their practical significance, directly reflecting the relative costs of FNs and FPs.

The Hessian, representing the second-order derivative of the standard binary cross-entropy loss, is:

$$h_i = \hat{y}_i(1 - \hat{y}_i). \qquad (9)$$

To reflect the class-dependent importance of FNs and FPs, we weight the Hessian as:

$$h_i = \hat{y}_i(1 - \hat{y}_i) \times (y_i \cdot w_{FN} + (1 - y_i) \cdot w_{FP}). \qquad (10)$$

This ensures the second-order derivative reflects the importance of each type of misclassification.

Furthermore, we selected a set of well-known ML algorithms and state-of-the-art methods as baselines for evaluating our proposed approach. Precisely, Multilayer Perceptron (MLP) captures complex patterns, KNN effectively utilizes distance metrics, SVM manages nonlinear decision boundaries, Linear Discriminant Analysis (LDA) enhances class separability through dimensionality reduction, XGB excels in structured data processing and performance boosting, and CNN leverages spatial hierarchies in data. For MLP, we employed a hidden-layer architecture with 100 neurons per layer, ReLU activation, the Adam optimizer, and 100 training epochs. KNN is configured with k=10, selected after evaluating values between 3 and 15 using Euclidean distance. We employed SVM with an Radial Basis Function (RBF) kernel and a regularization parameter C = 0.9. LDA is applied with the singular value decomposition (SVD) solver and a convergence tolerance of 0.0001, optimizing class separability. We also employed default XGB with a learning rate of 0.1, 100 boosting rounds, and a maximum tree depth of 6. For the CNN, we implemented a three-layer convolutional architecture with max-pooling and ReLU activations. The first layer applies 32 filters with a (3,3) kernel, followed by a second layer with 64 filters and a (3,1) kernel, both with max-pooling (2,1). After flattening, a fully connected layer with 128 neurons and ReLU activation captures complex patterns. The final output layer consists of two neurons with sigmoid activation for multi-label classification. The model is optimized using Adam (learning rate = 0.001) with binary cross-entropy loss and accuracy as the evaluation metrics. To mitigate overfitting, we utilized dropout layers with a 0.5 rate.

Regarding dataset splitting, we employed both Subject-Wise Cross-Validation (SWCV) and Leave-One-Out Cross-Validation (LOOCV). In SWCV, we partitioned each subject's data into 80% training and 20% test sets, ensuring that no windows overlap to prevent data repetition, which could otherwise reduce the number of samples. We train each model on the training set and evaluate it on the test set for each subject, then repeat this process for all five subjects. The final performance metric reflects the average results across all subjects. In LOOCV, we trained the model using data from four subjects and evaluated it on the remaining subject. We repeated this process five times, each time leaving out a different subject for evaluation. Additionally, the procedure was performed for each combination of time points and feature subsets, ensuring a comprehensive assessment of the model's performance across diverse individuals.

TABLE 3. Features of Hardware Platforms

Intel i9-13900	Raspberry Pi 5
Intel(R) Core (TM) i9-13900	Broadcom BCM2712
Max: 5.6 GHz, Min: 0.8 GHz	Up to 2.4 GHz
x86_64	ARM Cortex-A76
24 Cores	4 Cores
36 MiB Shared L3 Cache	2 MiB Shared L3 Cache
32 GB RAM	8 GB LPDDR4X

F. EVALUATION METRICS

To evaluate our classification models, we used several key metrics. Accuracy measures the ratio of correctly classified instances to the total number of instances. Precision reflects how many of the predicted positive instances are actually correct, while recall indicates how many of the actual positive instances were correctly identified. The F1-score combines precision and recall into a single metric by calculating their harmonic mean, offering a balanced assessment of classification performance. In multi-class classification, we use macro-averaged precision, recall, and F1-score to evaluate performance [37]. Accuracy offers an overall measure of correctness, while precision and recall focus on the quality of positive predictions. The F1-score provides a balanced assessment by combining precision and recall.

G. HARDWARE PLATFORMS

In addition to accuracy and performance metrics, we evaluate the computational efficiency of algorithms, which is critical for wearable EEG applications. Real-time predictions support continuous monitoring, while memory is limited and battery capacity must be preserved through energy-efficient computation. Processing time, memory usage, and energy consumption are key metrics. We measure feature extraction, model loading, and prediction on 15-second windows across two hardware platforms: (i) Intel i9-13900 (13th Gen Raptor Lake), representing a high-end edge device, and (ii) Raspberry Pi 5 (8 GB RAM), representing a lightweight edge device (Table 3). The platforms run Ubuntu 22.04.4 LTS and Raspberry Pi OS (Debian Bookworm), respectively. We measure processing time and memory usage in Python (v3.10.12) using `timeit` and `tracemalloc`. On the Intel platform, we measure core energy consumption using the Running Average Power Limit (RAPL) interface available on Intel processors [38], [39] with the `powercap` framework on Linux. On the Raspberry Pi platform, we utilize the command-line tool `vcgencmd` [40].

IV. RESULTS AND COMPARISON

This section presents our study's findings on both accuracy and computational efficiency, and provides a comparison with existing research.

A. RESULTS

Figure 3 illustrates the selection frequency of features, highlighting their relative importance. We assessed the perfor-

mance of various models using classification metrics, with a particular focus on the impact of feature selection on accuracy.

Figure 4 demonstrates the effect of varying the number of selected features from 21 to 30 on the accuracy of the XGB and cXGB models. The first 21 features are chosen based on their highest selection frequency in Figure 3, and additional features are incrementally included in order of selection frequency until all 30 are selected. The figure demonstrates that reducing the feature set by 10 results in only a 1% drop in accuracy for both models. Additionally, it highlights that the cXGB model achieves a more balanced distribution of FP and FN and higher accuracy than the XGB model.

Table 4 presents the mean and standard deviation of various performance metrics for the evaluated models. These values are computed across feature lengths ranging from 20 to 30 for all five subjects. The results indicate that different algorithms demonstrate varying sensitivity to the number of selected features, with models such as MLP being more affected than XGB and cXGB.

In the SWCV approach, training and testing are performed on data from the same individual without overlap. This setup yielded insufficient data for training a CNN; consequently, no results are reported for the CNN model. Among the evaluated algorithms, XGB and cXGB achieved the highest performance, with accuracies of 97.89% and 98.29%, respectively, in distinguishing focused states from other states.

Table 5 presents a performance comparison of various algorithms using LOOCV. Unlike the SWCV approach, the CNN model is included in this analysis because the LOOCV method provides sufficient data for effective training. The table shows that our cXGB model outperforms all other algorithms across F1-score, accuracy, recall, and precision. This superior performance is largely attributable to our adaptive learning rate strategy and the combination of FN and FP weights in gradient optimization, which enhances the model's ability. These targeted enhancements underscore our model's ability to generalize across subjects, thereby reinforcing its robustness and effectiveness across diverse evaluation settings.
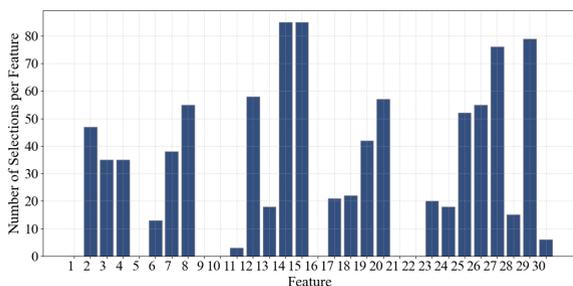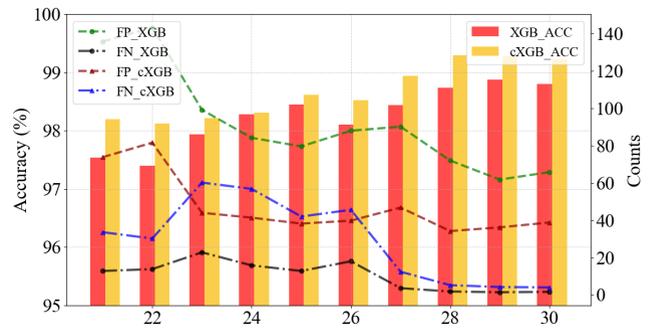
B. ACCURACY - COMPARISON

To ensure a fair and strict comparison, we selected the most recent studies in the literature that closely match the experimental conditions: all methods use the same dataset and apply equal window lengths for signal segmentation. According to existing publicly accessible EEG datasets, the dataset used in this study is the only one that provides clearly annotated attention labels suitable for benchmarking attention detection; therefore, our comparisons focus on studies that use the same dataset, along with six widely adopted machine learning algorithms evaluated under different cross-validation schemes. For precise evaluation, Table 6 compares the three-class classification performance of our proposed method cXGB that is trained on a subset of 20 optimally selected features against the works of Aci et al. [13] and Khare et al. [41]. The table presents subject-specific accuracy scores alongside the overall average accuracy across all subjects. Our method consistently outperforms both benchmarks in all cases except for subject 1 (SB1), where it performs slightly worse. The improvements range from 1% to 22% across subjects, with an average accuracy gain of 11.52% over the results presented in Khare et al.'s work and 2.53% over those reported by Aci et al.'s work. This highlights cXGB's robustness in handling inter-subject variability while maintaining superior classification performance.

The number of directly comparable studies on binary attention detection was extremely limited, with practically one work, namely Al-Nafjan et al. [23]. We employed 5-fold cross-validation on the entire dataset for this comparison. As demonstrated in Table 7, our cXGB method consistently outperforms the work of Al-Nafjan et al. across all evaluated metrics. Notably, our approach achieves a 2.76% higher accuracy. The improvement in recall is particularly significant: while Al-Nafjan et al.'s method shows low recall, our cXGB demonstrates a substantial 77.55% increase. It is noteworthy that none of the previous studies have reported their accuracy using LOOCV, despite its significance in assessing the generalizability of attention detection algorithms.

FIGURE 3. Frequency of Feature Selection Across 85 Iterations.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2026.3656851

Jahanjoo *et al.*: Preparation of Papers for IEEE TRANSACTIONS and JOURNALS

TABLE 4. Performance Metrics with Mean $\pm$ Standard Deviation for Different Binary Classification Models with SWCV

Metric	MLP	KNN	SVM	LDA	XGB	cXGB	CNN
F1-Score	59.20% $\pm$ 7.5	73.43% $\pm$ 5.0	61.18% $\pm$ 6.1	74.77% $\pm$ 3.7	97.90% $\pm$ 0.4	**98.31% $\pm$ 0.4**	$\times$
Accuracy	56.12% $\pm$ 3.1	73.94% $\pm$ 2.8	55.11% $\pm$ 3.1	75.93% $\pm$ 2.8	97.89% $\pm$ 0.4	**98.29% $\pm$ 0.4**	$\times$
Recall	56.12% $\pm$ 3.4	73.40% $\pm$ 2.6	55.11% $\pm$ 3.1	74.69% $\pm$ 3.3	97.93% $\pm$ 0.4	**98.31% $\pm$ 0.4**	$\times$
Precision	62.66% $\pm$ 8.5	73.54% $\pm$ 3.0	68.77% $\pm$ 9.5	75.21% $\pm$ 4.4	97.88% $\pm$ 0.4	**98.31% $\pm$ 0.4**	$\times$

$\times$ Not applicable due to insufficient data (see the text for further explanations)

TABLE 5. Performance Metrics with Mean $\pm$ Standard Deviation of Binary Classification Models with LOOCV

Metric	MLP	KNN	SVM	LDA	XGB	cXGB	CNN
F1	74.80% $\pm$ 0.50	70.08% $\pm$ 1.96	77.26% $\pm$ 0.18	73.18% $\pm$ 2.17	85.93% $\pm$ 2.15	90.97% $\pm$ 3.35	87.29% $\pm$ 2.9
ACC	78.53% $\pm$ 1.28	83.27% $\pm$ 1.42	78.54% $\pm$ 1.33	85.82% $\pm$ 1.66	82.23% $\pm$ 1.30	89.62% $\pm$ 2.23	77.79% $\pm$ 2.8
Recall	78.53% $\pm$ 1.28	72.32% $\pm$ 1.96	78.54% $\pm$ 1.33	76.00% $\pm$ 2.17	87.86% $\pm$ 3.13	99.95% $\pm$ 0.03	98.41% $\pm$ 0.69
Precision	71.42% $\pm$ 8.06	67.99% $\pm$ 1.50	76.09% $\pm$ 7.94	70.57% $\pm$ 1.01	81.00% $\pm$ 2.15	84.24% $\pm$ 3.35	78.78% $\pm$ 2.0

C. COMPUTATIONAL EFFICIENCY - CLASSIFICATION

Table 8 summarizes time, energy, and memory usage for feature extraction over a 15-second window on both platforms. On the Intel platform, using a reduced set of 20 features reduced processing time by 9.1%, current memory by 38.7%, peak memory by 8.3%, and energy consumption by 8.3%. Improvements are more pronounced on the Raspberry Pi: processing time is decreased by 15%, energy by 14%, and current and peak memory by 14% and 13%, respectively, likely due to the ARM processor's power-efficient design.

Figure 5 shows the prediction time measured over 16,935 overlapping windows of 15-second EEG signals for each classification model on the Raspberry Pi 5. On a per-window basis, most models complete prediction in well under 0.1 ms, with XGBoost requiring approximately 0.001 ms (binary) and 0.002 ms (ternary) per window. The slowest model, SVM, takes approximately 0.331 ms (binary) and 0.636 ms (ternary) per window. When including feature extraction time, these results indicate that real-time execution is fully feasible on the Raspberry Pi 5, even for the slowest model.

V. DISCUSSION

Our study introduces the cXGB algorithm and offers insights into the performance, personalization, generalizability, and efficiency of ML models for detecting focused, unfocused, and drowsy states using EEG data. We benchmarked cXGB against standard machine learning methods and state-of-the-art studies. Aci et al. [13] and Al-Nafjan et al. [23] used classical classifiers with multi-channel EEG but did not pro-

TABLE 7. Performance Comparison of Binary Classification Models with 5-Fold Cross-Validation

	Accuracy	Recall	Precision	F-measure
Al-Nafjan [23]	87%	14%	83%	77%
Proposed method	**89.76%**	**91.55%**	**89.76%**	**90.93%**

vide justification for channel selection and lacked subject-independent evaluation. Their results may therefore be overfit to individual users. In contrast, we chose the AF3 channel because it showed apparent differences between attention states and is located in the prefrontal region, which is linked to attention. We used RFE to select the most important features, helping reduce noise and focus on meaningful patterns. On the other hand, Khare et al. [41] used a complex ensemble of wavelet-based methods to process seven-channel EEG, but their system demands high computational power. Additionally, we evaluated a baseline CNN model to compare its accuracy, prediction time, energy consumption, and memory usage with our proposed method. CNN computational footprint was almost always higher than that of the proposed cXGB approach. Given our design objective of maximizing accuracy under strict efficiency constraints for wearable deployment, we did not extend the comparison to more advanced deep architectures such as attention-based networks (transformers) or graph neural networks. These models, although powerful in large-scale EEG decoding tasks, typ-

TABLE 6. Accuracy Comparison Across Subjects for Three-Class Classification

	SB1	SB2	SB3	SB4	SB5	Average
Aci [13]	**96.70%**	89.70%	88.60%	93.50%	90.10%	91.72%
K.Khare [41]	87.45%	84.61%	84.12%	81.47%	75.98%	82.73%
Proposed method	92.69%	**92.32%**	**94.16%**	**94.38%**	**97.70%**	**94.25%**

TABLE 8. Time, Core Energy, and Memory & Peak Memory Usage for Feature Extraction per 15-second Window (Mean $\pm$ Standard Deviation)

Metric	Intel(R) i9-13900	Raspberry Pi 5
Time (ms)	8.77 $\pm$ 0.44	20.487 $\pm$ 0.023
Core Energy (mJ)	135.2 $\pm$ 1.7	63.0 $\pm$ 0.7
Memory Usage (B)	9280 $\pm$ 1	9332 $\pm$ 0.2
Peak Memory Usage (B)	110216 $\pm$ 0	102068 $\pm$ 0

ically incur even greater computational costs than CNNs, making them less suitable for edge platforms with limited resources. By focusing on a lightweight yet accurate solution, our work aligns more closely with the practical requirements of portable EEG devices. Utilizing a single EEG channel substantially reduces computational requirements. The proposed cXGB algorithm achieves real-time performance for 15-second windows on the Raspberry Pi. This efficiency makes our algorithm particularly suitable for applications involving lightweight, portable EEG devices designed for everyday use, especially in scenarios utilizing commercial EEG headbands with limited channels. Although LDA is the most efficient algorithm, the proposed cXGB algorithm achieves significantly higher accuracy, outperforming LDA by over 20% in the SWCV experiment and by 4% in the LOOCV experiment. Detecting attention in everyday settings, such as driving, office work, or learning, can reduce fatigue-related risks, boost productivity, and support cognitive monitoring. This has clinical relevance for the management of attention deficits, drowsiness, and cognitive decline. Our findings underscore the potential of single-channel EEG systems, offering a more convenient and faster solution without compromising performance for such applications.

To the best of our knowledge, no publicly available datasets provide real-world EEG data collected from wearable devices in everyday, uncontrolled environments. Consequently, model performance in actual use cases may differ from lab-based results. In the absence of such datasets, lab-collected recordings remain the most practical option for benchmarking. To evaluate generalizability, we adopted LOOCV, which tests the model on entirely unseen subjects. This subject-independent validation reduces bias and offers a more realistic estimate of how well the algorithm might perform across different individuals.

VI. CONCLUSION

In this study, we proposed the cXGB model trained on a carefully selected set of features using single-channel EEG data. Our approach outperformed existing methods in detecting focused, unfocused, and drowsy states, achieving 98.29% accuracy in binary classification and 94.25% in three-class

classification, which demonstrates the individualization capability of our method. The model also demonstrated strong generalization, achieving 89.62% accuracy when applied to previously unseen individuals. To assess practical deployment, we implemented the model on a resource-constrained edge device (Raspberry Pi 5). Feature reduction led to substantial efficiency gains: processing speed improved by 15%, energy consumption decreased by 14%, and memory usage was reduced by 14%. Each 15-second EEG window was processed in real-time, with inference latency ranging from 35 to 60 ms.

The model's high accuracy and computational efficiency, achieved using a *single* EEG channel, highlight its potential for portable, personalized neurotechnology. These results support integrating lightweight, individualized AI models into wearable systems for real-time attention monitoring, advancing the broader vision of precision health.

In future work, we plan to evaluate the model using real-world data collected during daily work routines to assess attention patterns and explore strategies for improving focus and productivity.

REFERENCES

[1] N. TaheriNejad and D. Pollreisz, "Assessment of physiological signals during happiness, sadness, pain or anger," in *6th MobiHealth*, 2017.

[2] Y. Gao, A. Liu, L. Wang, R. Qian, and X. Chen, "A self-interpretable deep learning model for seizure prediction using a multi-scale prototypical part network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1847–1856, 2023.

[3] D. Freismuth and N. TaheriNejad, "On the treatment and diagnosis of attention deficit hyperactivity disorder with eeg assistance," *Electronics Journal*, pp. 1–18, 2022.

[4] U. Maria Muna, M. Mehedi Hasan Shawon, M. Jobayer, S. Akter, and S. Rahman Sabuj, "Sstaf: Spatial-spectral-temporal attention fusion transformer for motor imagery classification," *IEEE Access*, vol. 13, pp. 158 855–158 869, 2025.

[5] C. Li, P. Li, Y. Zhang *et al.*, "Effective emotion recognition by learning discriminative graph topologies in eeg brain networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[6] L. Gong, W. Chen, and D. Zhang, "An attention-based multi-domain bi-hemisphere discrepancy feature fusion model for eeg emotion recognition," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 10, pp. 5890–5903, 2024.

[7] P. Saha, A. K. A. Kunju, M. E. Majid *et al.*, "Novel multimodal emotion detection method using electroencephalogram and electrocardiogram signals," *Biomedical Signal Processing and Control*, vol. 92, p. 106002, 2024.

[8] P. Traunmuller, A. Jahanjoo, S. Khooyooz *et al.*, "Wearable healthcare devices for monitoring stress and attention level in workplace environments," *arXiv preprint arXiv:2406.05813*, 2024.

[9] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability," *Sensors*, vol. 17, no. 9, p. 1991, 2017.

[10] Z. Guo, Y. Pan, G. Zhao, S. Cao, and J. Zhang, "Detection of driver vigilance level using eeg signals and driving contexts," *IEEE Transactions on Reliability*, vol. 67, no. 1, pp. 370–380, 2017.

[11] P. Abhang, B. Gawali, and S. Mehrotra, *Technical Aspects of Brain Rhythms and Speech Parameters*. Academic Press, 2016, ch. 3, pp. 51–79.

[12] P. Kaushik, A. Moye, M. v. Vugt, and P. P. Roy, "Decoding the cognitive states of attention and distraction in a real-life setting using eeg," *Scientific Reports*, vol. 12, no. 1, p. 20649, 2022.

[13] Çiğdem İnan Acı, M. Kaya, and Y. Mishchenko, "Distinguishing mental attention states of humans via an eeg-based passive bci using machine learning methods," *Expert Systems with Applications*, vol. 134, pp. 153–166, 2019.

[14] Z. Wang, Y. Ouyang, and H. Zeng, "Arfn: An attention-based recurrent fuzzy network for eeg mental workload assessment," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–14, 2024.
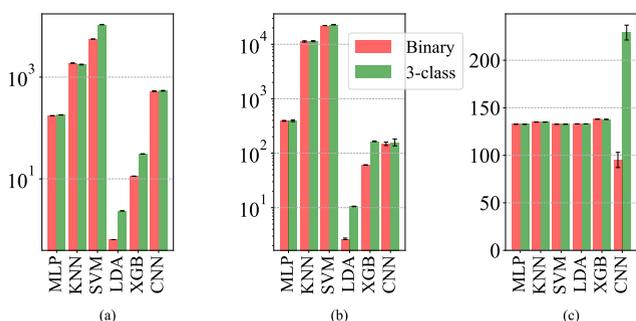
FIGURE 5. Prediction (a) Time (ms), (b) Core Energy Consumption (mJ), and (c) Memory Usage (KiB) across Classification Models on Raspberry Pi 5.

IEEE *Access*

[15] Y. Zhou, J. Jiang, L. Wang, S. Liang, and H. Liu, "Enhanced cognitive load detection in air traffic control operators using eeg and a hybrid deep learning approach," *IEEE Access*, vol. 13, pp. 12 127–12 137, 2025.

[16] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of neural engineering*, vol. 16, no. 5, p. 051001, 2019.

[17] H. Zhang and H. Li, "Transformer-based eeg decoding: A survey," *arXiv preprint arXiv:2507.02320*, 2025.

[18] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[19] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[20] H. Asemi and N. Farajzadeh, "Improving eeg signal-based emotion recognition using a hybrid gwo-xgboost feature selection method," *Biomedical Signal Processing and Control*, vol. 99, p. 106795, 2025.

[21] F. Wang, Y.-C. Tian, X. Zhang, and F. Hu, "An ensemble of xgboost models for detecting disorders of consciousness in brain injuries through eeg connectivity," *Expert Systems with Applications*, vol. 198, p. 116778, 2022.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[23] A. Al-Nafjan and M. Aldayel, "Predict students' attention in online learning using eeg data," *Sustainability*, vol. 14, no. 11, 2022. [Online]. Available: https://www.mdpi.com/2071-1050/14/11/6553

[24] N. G. Seshadri, S. Agrawal, B. K. Singh, B. Geethanjali, V. Mahesh, and R. B. Pachori, "Eeg based classification of children with learning disabilities using shallow and deep neural network," *Biomedical Signal Processing and Control*, vol. 82, p. 104553, 2023.

[25] H. Yang, J. Wu, Z. Hu, and C. Lv, "Real-time driver cognitive workload recognition: Attention-enabled learning with multimodal information fusion," *IEEE Transactions on Industrial Electronics*, vol. 71, no. 5, pp. 4999–5009, 2023.

[26] C. Y. Jin, J. P. Borst, and M. K. Van Vugt, "Predicting task-general mind-wandering with eeg," *Cognitive, Affective, & Behavioral Neuroscience*, vol. 19, pp. 1059–1073, 2019.

[27] Z. Mohamed, M. El Halaby, T. Said, D. Shawky, and A. Badawi, "Characterizing focused attention and working memory using eeg," *Sensors*, vol. 18, no. 11, p. 3743, 2018.

[28] T. Suhail, K. Indiradevi, E. Suhara, S. A. Poovathinal, and A. Ayyappan, "Distinguishing cognitive states using electroencephalography local activation and functional connectivity patterns," *Biomedical Signal Processing and Control*, vol. 77, p. 103742, 2022.

[29] D. Devi and S. Sophia, "Ga-cnn: analyzing student's cognitive skills with eeg data using a hybrid deep learning approach," *Biomedical Signal Processing and Control*, vol. 90, p. 105888, 2024.

[30] E. Magosso, G. Ricci, and M. Ursino, "Modulation of brain alpha rhythm and heart rate variability by attention-related mechanisms," *AIMS neuroscience*, vol. 6, no. 1, p. 1, 2019.

[31] NeuroSky, "Why locate the sensor at fp1," http://support.neurosky.com/kb/science/why-locate-the-sensor-at-fp1, [Online; Accessed: Mar. 1, 2025].

[32] Muse, "Science," https://choosemuse.com/pages/science, [Online; Accessed: Mar. 1, 2025].

[33] Macrotellect, "Brainlink usage," https://macrotellect.tawk.help/article/faqs-on-brainlink-usage, [Online; Accessed: Mar. 1, 2025].

[34] E. S. Kappenman and S. J. Luck, "The effects of electrode impedance on data quality and statistical significance in erp recordings," *Psychophysiology*, vol. 47, no. 5, pp. 888–904, 2010.

[35] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel *et al.*, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.*, vol. 12, no. null, p. 2825–2830, nov 2011.

[36] I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," *arXiv preprint arXiv:1608.03983*, 2016.

[37] M. Grandini, E. Bagli, and G. Visani, "Metrics for multi-class classification: an overview," *arXiv preprint arXiv:2008.05756*, 2020.

[38] Intel Corporation, *Intel 64 and IA-32 Software Developer Manuals - Volume 3*, https://www.intel.com/content/www/us/en/developer/articles/technical/intel-sdm.html, 2025, [visited on 05.03.2025].

[39] K. N. Khan, M. Hirki, T. Niemi, J. K. Nurminen, and Z. Ou, "Rapl in action: Experiences in using rapl for power measurements," vol. 3, no. 2, mar 2018. [Online]. Available: https://doi.org/10.1145/3177754

[40] Raspberry Pi Foundation, "Power Supplies Documentation," https://github.com/raspberrypi/documentation/blob/develop/documentation/asciidoc/computers/raspberry-pi/power-supplies.adoc, 2025, [visited on 05.03.2025].

[41] S. K. Khare, V. Bajaj, N. B. Gaikwad, and G. R. Sinha, "Ensemble wavelet decomposition-based detection of mental states using electroencephalography signals," *Sensors*, vol. 23, no. 18, p. 7860, 2023.

ANICE JAHANJOO received her B.Sc. degree in Computer Software Engineering and her M.Sc. degree in Artificial Intelligence and Robotics from Shahid Chamran University of Ahvaz, Iran, in 2015 and 2018, respectively. She is currently pursuing a Ph.D. degree at Vienna University of Technology, Vienna, Austria. Her research interests include machine learning, cognitive recognition, biomedical signal processing, and wearable healthcare technology.

VIMALA BAUER was born in Mauritius. She received a B.Eng. degree in Computer and Microelectronic Systems from the University of Liverpool, UK, and M.S. degrees in Computer and Information Science from the University of Oregon, USA, and in Computer Engineering from the University of Heidelberg, Germany, in 2023. She is currently pursuing a Ph.D. in Computer Engineering at the University of Heidelberg, focusing on designing efficient hardware for wearable healthcare devices.

She has been working as a freelance lecturer and has developed distance learning and laboratory materials for three Universities of Applied Sciences in Germany. Her teaching interests include computer architecture and embedded systems programming.

She is a recipient of the National Mauritius State Scholarship, the Fulbright Scholarship, and the Morris Ranger Undergraduate Award.

SOHEIL KHOOYOOZ received his B.Sc. degree in Biomedical Engineering (Bioelectrics) from Amirkabir University of Technology, Tehran, Iran, in 2020, and his M.Sc. degree in Biomedical Engineering (Bioelectrics) from Sharif University of Technology, Tehran, Iran, in 2023. He is currently a Ph.D. candidate in the Department of Computer Engineering at Heidelberg University, Heidelberg, Germany. His research interests include wearable healthcare systems, biomedical signal processing, and the application of AI/ML to these domains.

IEEE *Access*

MOSTAFA HAGHI (Member, IEEE) received the Ph.D. degree from the Center for Life Science Automation, the University of Rostock, Rostock, Germany, in 2019. He is currently a Postdoctoral Researcher with the Computer Architecture Team, at the Institute for Computer Engineering, Heidelberg University, Germany. His major research interests include wearable healthcare, hardware/software-embedded systems, bioelectronics, and sleep medicine. He is the Co-Chair of the Accident and Emergency Informatics Working Group of the International Medical Informatics Association.

NIMA TAHERINEJAD received his Ph.D. degree in electrical and computer engineering from The University of British Columbia (UBC), Vancouver, Canada, in 2015. He is currently a full professor at Heidelberg University, Heidelberg, Germany and affiliated with TU Wien (formerly known also as Vienna University of Technology), Vienna, Austria. His areas of work include computer architecture (especially memory-centric and approximate computing), cyber-physical and embedded systems, systems on chip, memristor-based circuit and systems, self-* systems, and smart health-care. He has published three books, five patents, and more than 100 articles. Prof. Taherinejad has served as an editor of many journals, an organizer and a chair of various conferences and workshops. He has received several awards and scholarships from universities, conferences, and competitions he has attended. This includes the Best University Booth award at DATE 2021, First prize in the 15th Digilent Design Contest (2019) and in the Open-Source Hardware Competition at Eurolab4HPC (2019) as well as Best Teacher and Best Course awards at TU Wien (2020).

. . .