

Non-Intrusive Cognitive Load Estimation Using Footstep-Induced Structural Vibration Signals

MAINAK CHAKRABORTY, CHANDAN,
BODHIBRATA MUKHOPADHYAY and SUBRAT KAR

ABSTRACT

Cognitive load estimation is crucial for understanding human performance during demanding tasks. Elevated cognitive load can impair decision-making and increase the risk of fatigue or burnout. Consequently, real-time monitoring of cognitive load is essential in high-stakes environments such as healthcare, where even subtle changes in human movement can have significant consequences. Conventional methods for cognitive load assessment predominantly rely on neuro-imaging, physiological signals, or subjective reporting, which are often not real-time, intrusive and can cause user discomfort. In this study, we propose the use of structural vibrations as a non-intrusive and privacy-preserving modality for cognitive load estimation, offering a promising alternative in healthcare and other sensitive domains. We conducted controlled experiments with eight participants under two walking conditions: a baseline of normal walking and a dual-task condition where participants walked while engaging in cognitively demanding interactions with a conversational agent. Comparative analysis revealed distinct changes in gait dynamics and footstep-induced structural vibration patterns under cognitive load. Our experimental results identified clear sensor signatures corresponding to varying levels of cognitive demand, demonstrating the potential of structural vibration metrics as indicators of cognitive load. We achieved 68.00% accuracy in cognitive load classification using structural vibration signals. Furthermore, for person identification, we attained an accuracy of 82.00% based on structural vibration data.

INTRODUCTION

Cognitive load, defined as the mental effort exerted by an individual's working memory during task performance, is a critical factor in understanding the efficiency and limitations of human information processing [1]. Cognitive load itself refers to the mental effort drawn from working memory during engagement with a cognitive activity or system. Accurate estimation requires recognition of its three primary types. *Intrinsic load* is

Mainak Chakraborty, PhD Student, Email: idz218413@iddc.iitd.ac.in. Centre for Sensors, Instrumentation and Cyber Physical System Engineering (SeNSE), Indian Institute of Technology Delhi (IIT Delhi), New Delhi, India

associated with the inherent difficulty of the task and is not modifiable by the researcher. *Extraneous load* arises from the manner in which information or tasks are presented, such as interface design, which can be minimized through careful design. *Germane load* represents the effort devoted to processing and learning from the task [2, 3]. Research demonstrates that cognitive load assessment is widely applied across psychology, neuroscience, and visual computing. To investigate cognitive load, researchers use variety of techniques, including electroencephalography (EEG), electro-dermal activity (EDA), eye-tracking, audio-speech recognition, heart rate variability (HRV), and accelerometers [4–8].

Electroencephalography (EEG) measures electrical brain activity using scalp electrodes and is notable for its high temporal resolution, making it well-suited for capturing rapid changes in cognitive states and assessing instantaneous cognitive load [6]. Electro-dermal activity (EDA) and heart rate variability (HRV) are intrusive measures that capture autonomic nervous system responses to cognitive demands. Recent studies have investigated these markers in participants observed significant changes when exposed to demanding mental tasks [7]. Subjective measures such as the NASA Task Load Index (NASA-TLX) and Workload Profile (WP) requires self-reported data that interrupts task flow [9].

Among non-intrusive methods, eye tracking features such as pupil dilation, eye movement patterns, and blink rate-has emerged as a valuable tool for assessing cognitive load [10, 11]. The analysis of speech patterns and linguistic features offers a promising approach to cognitive load assessment by leveraging natural communication data [5]. Research on management teams found that under high cognitive load, team members spoke more, used longer sentences, employed more disagreement words, increased their use of plural personal pronouns, and decreased their use of singular pronouns [12]. Research also suggests that cognitive load effects on gait can be measured using off-the-shelf accelerometers [8, 13]. Traditional cognitive load estimation methods, while valuable, often face limitations related to cost, invasiveness, and real-time applicability. Neuroimaging techniques, for instance, can be expensive and require specialized infrastructure. Physiological measures may be influenced by non-cognitive factors, and subjective reports are typically administered post-task and rely on introspection.

Our aim is to understand cognitive load in a more naturalistic setting, based on human gait. Early research in gait analysis has established a clear link between cognitive load and changes in gait parameters such as speed, stride length, and variability [13–15]. Structural vibration-based analysis, utilizing sensors such as geophones, presents a promising alternative with several advantages [16–20]. Compared to expensive neuroimaging equipment, geophones offer a cost-effective alternative that can be deployed in a wide range of settings without the need for specialized infrastructure, while providing real-time updates. Their non-invasive and privacy-preserving nature enables seamless integration into everyday environments, enhancing user comfort and ease of use.

The main contributions of this paper are as follows:

- We propose a novel framework for non-intrusive cognitive load estimation that

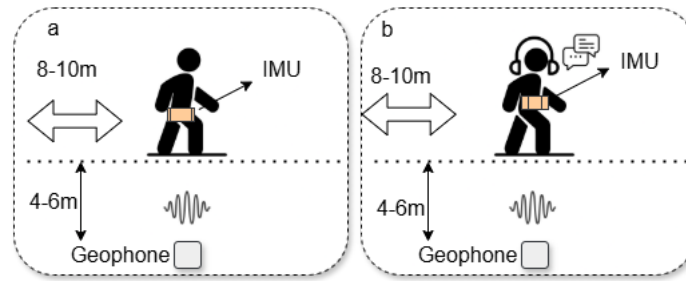


Figure 1: Experimental setup for gait recording using geophone and IMU sensors. (a) Baseline condition: participants walk naturally over an 8–10 meter path while IMU sensor and vibration signals are recorded by a geophone placed 4–6 meters away. (b) Cognitive load condition: participants walk the same path while simultaneously engaging in a conversation with a custom Large Language Model (LLM), inducing cognitive distraction during walking.

leverages structural vibration sensors. Additionally, we use inertial measurement unit (IMU) sensors, which are attached to the body, to compare the findings. To the best of our knowledge, this is the first cognitive load estimation system based on footstep-induced structural vibrations.

- We conduct a comprehensive investigation of various gait features, capturing both person-variant and person-invariant characteristics relevant to cognitive load assessment.
- We validate our approach through real-world experiments, collecting data from eight individuals across two distinct conditions: (a) normal walking and (b) walking under cognitive load. Data were acquired using both sensor modalities, resulting in a total of 160 minutes of recorded structural vibration and IMU sensor data.

EXPERIMENTAL DESIGN

The primary aim of this study is to investigate the impact of increased cognitive load and to quantitatively assess its effects through measurements of structural vibration. A major challenge in this work was the lack of an existing dataset suitable for the research objectives. To address this, we curated a new dataset involving eight individuals across two experimental protocols. Our goal was to design a system capable of detecting cognitive overload by monitoring subtle changes in human movement, with the broader objective of supporting tasks that maximize germane cognitive load (the effort devoted to processing and meaningful learning), while managing intrinsic load (the inherent difficulty of the task) and minimizing extraneous load (the unnecessary complexity in the way information is presented). To simulate naturalistic cognitive load, we employed a conversational agent, as walking and talking represent a common, ecologically valid dual-task scenario.

As shown in figure 1, the experimental design was straightforward: participants were

GPT: Hey there! I'm a GPT trained to give you multiple cognitive challenging tasks. I'll give you quick brain teasers while you're out and about, adjusting the challenge as we go. If you don't know the answer simply say no, I will ask the next question. Ready for your first one?

User: Ready!

GPT: Great. Let's try a memory one. Listen closely: *Maroon, Blue, Green, Yellow, Red*. Which color was before Yellow?

User: Green.

GPT: Great. Listen and remember this sequence: One-A, Two-B, Three-C. Can you repeat that sequence?

User: One-A, Two-B, Three-C.

GPT: Perfect! You got it. Now, based on that pattern, what number is associated with E?

....

Figure 2: A sample conversation with the GPT and a user while walking

instructed to walk along a predefined 8-10m trajectory under two distinct conditions. In the baseline condition, each participant walked for five minutes without any external distractions, ensuring that no other individuals were present in the environment to minimize confounding variables and establish a reference dataset.

In the second condition, we instructed participants to walk for five minutes while simultaneously engaging in a conversation with a custom Large Language Model (LLM) specifically fine-tuned for this scenario [21]. The LLM dynamically adjusted the difficulty of short cognitive tasks—such as questions, riddles, and one/two-back tests—based on the participant's performance. Our intuition was that, compared to normal walking, a significant distraction would cause noticeable shifts in gait patterns, which we could capture using geophone and IMU sensors. Figure 2 illustrates an example of a conversation with the LLM. We informed participants that responding to the LLM's queries was the primary task, while walking was secondary; they were allowed to pause, think, and adjust their walking to better facilitate their responses. As in the baseline condition, we controlled the environment to ensure that no other individuals or activities were present during the experiment. Through this protocol, we aimed to study the effects of cognitive load, induced by the conversational task, on gait and the resulting structural vibrations.

EXPERIMENTAL SETUP

In this study, we incorporated two types of sensors, namely geophone and Inertial Measurement Units (IMUs), as shown in figure 3¹. A comprehensive explanation of these sensors is provided below:

¹The data collection process adhered to the by the Institutional Ethics Committee(IEC). Ethical approval from the IEC, Indian Institute of Technology Delhi, was obtained (Ref. number IITD/IEC/P-059)

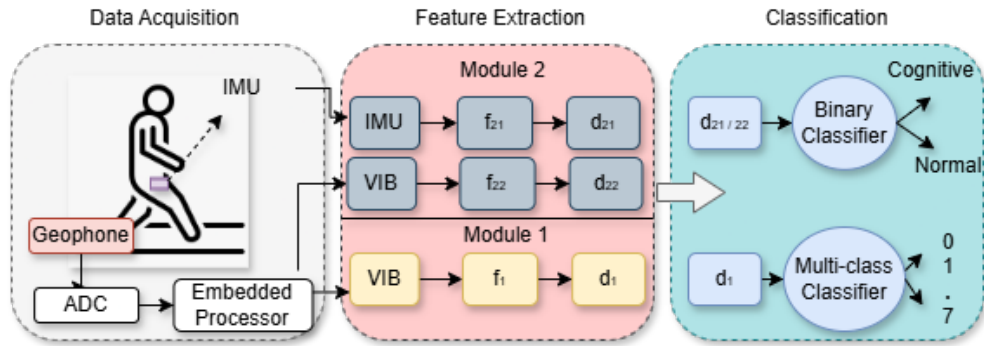


Figure 3: Overview of the sensor-based cognitive load estimation framework. Geophone and IMU sensors capture subtle movement dynamics during walking tasks. IMU and Structural Vibration (VIB) signals are processed through two feature extraction f_{21} and f_{22} modules. Module 1 extracts vibration features for multi-class person identification, while Module 2 is used for binary classification of cognitive states.

- **Geophone** : We used one geophone sensor with a sensitivity of 2.88 V/m/sec, and a gain of 10 is used for both experiments. Vibration signals were recorded using a Logic sound card hat equipped with a 16-bit analog-to-digital converter (ADC) operating at a sampling rate of 8 kHz. The sound card interfaced with a Raspberry Pi 5, having 8 GB of RAM and 32 GB of storage. The geophone sensor operates passively, requiring no external power supply, and draws minimal operating current (measured as 5–10 mA at 5 V using the Nordic Power Profiler v2 kit).
- **Inertial Measurement Units (IMU):** We have used off-the-shelf smartphone based IMU sensors to record linear acceleration data in three direction (x,y,z). The average current consumption is approximately 0.24 mA for the accelerometer. The resolution of the accelerometer is 0.004788 m/s². The minimum achievable delay between successive sensor readings is 2000 microseconds. Data acquisition was performed using MATLAB Online (version R2022b) [22], with the smartphone positioned securely in the participant’s front trouser pocket to simulate typical real-world carrying scenarios. The acquisition system was configured with a sampling frequency of 10 Hz to maintain a consistent temporal resolution across all datasets. For each participant, continuous inertial data were recorded over a five-minute session under experimental conditions.

DATA PRE-PROCESSING AND FEATURE EXTRACTION

As shown in table I, our system mainly consists of two modules: 1) Person identification from footsteps, and 2) Cognitive load estimation using sensors. In data pre-processing, for acceleration data, the tri-axial time-series signals are segmented into fixed-length, non-overlapping windows of 500ms ($L = 5$) samples. For structural vibration data, a one-dimensional time-series signal is segmented into non-overlapping windows of a

TABLE I: Feature Summary

Features
<i>Module 1: Person Identification (f_1)</i>
GMM Inputs: Skew., Kurt., Freq. (40-160 Hz) CWT Output: Time-Freq. Map
<i>Module 2: Cognitive Load (f_{2x})</i>
Accelerometer Signals (f_{21})
Time: μ , σ , RMS, PTP, SMA, ZCR, Med. Freq: H , C_{spec} , f_{dom}
Structural Vibration signals (f_{22})
Time: RMS, μ , σ , γ_1 , γ_2 , PTP, E , ZCR Peak: N_{peaks} , μ_{peak} , σ_{peak} , f_{step} Freq: f_{dom} , M_{dom} , C_{spec} , H_{spec} DWT E.: E_{approx} , $E_{detail1-4}$

fixed size of 250ms; the signal is sampled at a frequency $f_s = 8000$.

Module 1: Person identification from footsteps-induced structural vibration

In this study, we employ an unsupervised method to detect individual footstep events in vibration recordings [20]. We use a Gaussian Mixture Model (GMM) with two clusters, leveraging features such as skewness, kurtosis, and frequency-based energy bands extracted from 350 ms segments smoothed with a Gaussian window. After training the GMM on vibration data from a single subject, we extract footstep events, each represented by 1500 data points. We then converted these events to time-frequency images using the Continuous Wavelet Transform (CWT) with Morlet wavelets across 0–256 scales.

Module 2: Cognitive load estimation using sensors In this study we use distinct feature extraction methodologies tailored to the characteristics of accelerometer and structural vibration signals. For accelerometer signals, we calculate the mean (μ), standard deviation (σ), root mean square (RMS), peak-to-peak amplitude (PTP), signal magnitude area (SMA), zero-crossing rate (ZCR), and median. To capture the spectral properties of the accelerometer signals, frequency-domain features such as entropy (H), spectral centroid (C_{spec}), and dominant frequency (f_{dom}) are extracted from the Fast Fourier Transform (FFT). We compute these features across all three linear acceleration x, y and z.

Structural vibration signals encompass a broader range of time-domain features: RMS, mean (μ), standard deviation (σ), skewness (γ_1), kurtosis (γ_2), peak-to-peak amplitude (PTP), signal energy (E), and zero-crossing rate (ZCR). Additionally, peak detection techniques are utilized to derive features related to the signal’s extrema, including the number of peaks (N_{peaks}), the average (μ_{peak}) and standard deviation (σ_{peak}) of peak heights, and the average step frequency (f_{step}). Frequency-domain analysis of the structural vibration windows yields the dominant frequency (f_{dom}) and its magnitude (M_{dom}), the spectral centroid (C_{spec}), and the spectral entropy (H_{spec}). Furthermore, a multi-level discrete wavelet transform is applied to extract energy features from the approximation (E_{approx}) and detail coefficients ($E_{detail1}$, $E_{detail2}$, $E_{detail3}$, $E_{detail4}$), capturing time-frequency characteristics. This dual approach ensures a comprehensive representation

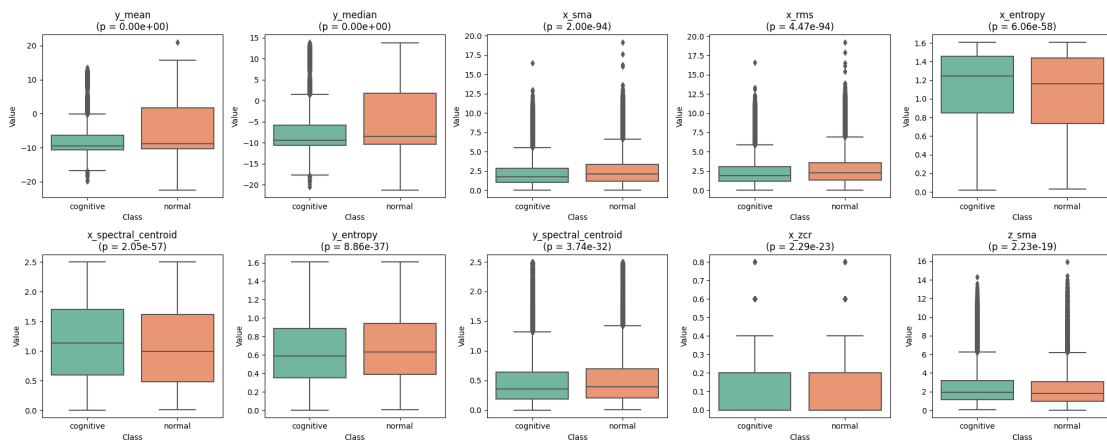


Figure 4: Top ten global accelerometer features differentiating cognitive from normal walking. Vertical motion (y-axis) is significantly reduced under cognitive load. Other key differences include signal energy, entropy, and frequency features across axes, revealing clear biomechanical adaptations.

of the information contained within both accelerometer and structural vibration datasets.

QUALITATIVE STUDY AND DATA VISUALIZATION

We analyzed IMU and structural vibration data to compare normal walking with walking under cognitive load. To identify the most discriminative features between the two experimental conditions, we used an analysis of variance (ANOVA)-based feature selection approach, followed by visualization via boxplots.

As shown in figure 4, participants showed significantly reduced vertical motion under cognitive load. The y-axis mean and median values were more negative, indicating a flatter walking pattern. Movement energy, measured by signal magnitude area (SMA) and root mean square (RMS), decreased, especially in the x-axis, suggesting reduced intensity. Spectral centroid values shifted, reflecting altered movement frequency. Along the anterior-posterior axis (X-axis), movement energy decreased, reflecting more controlled forward motion. Vertical displacement in the Y-axis was significantly reduced, indicating changes in gait mechanics. Furthermore, in the mediolateral direction (Z-axis), both entropy and signal magnitude decreased substantially, suggesting diminished side-to-side stability.

As shown in figure 5, RMS and standard deviation dropped by 50% under cognitive load, showing lower vibration magnitude. Peak-to-peak amplitude and total energy also declined, indicating more controlled foot impacts. Wavelet energies and dominant magnitude decreased across bands, confirming lower force transmission. The zero-crossing rate increased, reflecting more frequent, smaller steps. Kurtosis shifted, indicating changes in signal distribution.

Together, the IMU and vibration data paint a coherent picture: when performing a

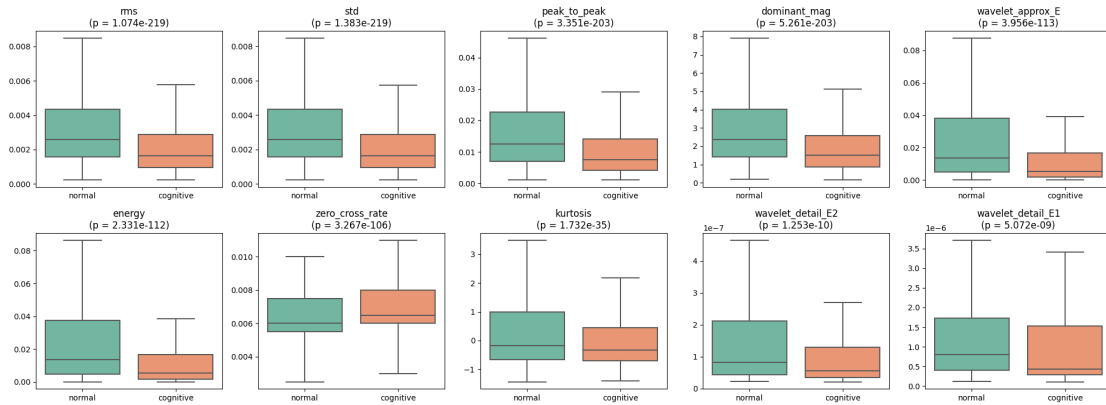


Figure 5: Top ten structural vibration features distinguishing normal vs. cognitive walking. Amplitude and energy features are lower during cognitive walking, indicating reduced force and more cautious steps. Zero-crossing rate is higher, reflecting more frequent, smaller movements. All features show highly significant differences.

TABLE II: Classification accuracies (in %) with standard deviations (in parentheses) for different sensing modalities using Random Forest (RF) and Support Vector Machine with RBF kernel (SVM-RBF).

Modality	Random Forest		Support Vector Machine	
	Accuracy	F1-score	Accuracy	F1-score
Acceleration	74.86 (1.25)	74.00 (0.40)	75.39 (0.88)	75.20 (0.45)
Structural Vibration	68.49 (1.50)	67.40 (1.00)	66.02 (0.54)	65.80 (0.20)

cognitive task while walking, individuals unconsciously shift to a more stable, energy-efficient gait characterized by reduced vertical displacement, less movement energy, altered movement frequency, more frequent but smaller steps, and diminished foot impact forces. These adaptations appear to help individuals maintain postural control when cognitive resources are divided. The robustness and consistency of these findings, especially the statistical significance across numerous features, provide strong evidence that cognitive loading elicits measurable, reproducible changes in gait biomechanics, that can be recorded via structural vibration and IMU sensors.

EXPERIMENTAL RESULTS AND DISCUSSION

We evaluated our system’s performance for binary classification using two different models: a Random Forest classifier and a Support Vector Machine (SVM) with an RBF kernel. We trained and validated both models using stratified five-fold cross-validation. The Random Forest classifier was configured with a maximum tree depth of 20. The SVM classifier utilized an RBF kernel with a regularization parameter $C=10$. Model performance was assessed based on standard metrics including accuracy and F1-score.



Figure 6: Training and testing loss (left) and accuracy (right) curves across 25 epochs.

Our analysis revealed a notable dependency of classification performance on sensor modality. It is evident that changes induced by cognitive load manifest more profoundly when the sensor is directly attached to the body, likely capturing subtle physiological or movement variations. Classification accuracy based on structural vibrations, where the sensor is not directly affixed to the body, achieved an accuracy of 68%. This result is considered impressive given the indirect nature of the measurement modality, demonstrating the potential of structural vibration analysis for detecting cognitive states. In comparison, IMU-based classification achieved a higher accuracy, benefiting from direct contact and reduced signal attenuation. Nevertheless, the ability of geophone-based structural vibration sensing to infer cognitive states highlights its promise for non-intrusive, scalable deployment in real-world environments where direct sensor attachment may not be feasible.

For person identification, we trained a deep learning classification model using a transfer learning strategy. We organized the dataset into a folder structure compatible with the PyTorch and split it into 80% training and 20% testing subsets. We applied basic preprocessing by converting training images to tensors and resizing testing images to 224×224 pixels before tensor conversion. We initialized a ResNet-18 model pretrained on ImageNet and replaced its final fully connected layer to match the number of target classes [23]. We fine-tuned all model parameters rather than freezing the backbone layers. During training, we employed the Adam optimizer with weight decay and used the Cross-Entropy Loss function to supervise learning. We adapted the learning rate dynamically using a ReduceLROnPlateau scheduler, which lowered the learning rate when the validation loss plateaued. The model is trained for 15 epochs, with batch size 16. As shown in the Figures, the trained model achieves an accuracy of 82% (1.00). As shown in the figure 6, the model demonstrates stable convergence, with test loss decreasing and test accuracy improving steadily, and performance stabilizes around 80%.

CONCLUSION

In this paper, we proposed a cognitive load estimation framework utilizing structural vibration and IMU sensors. One of the primary challenges in developing this system was the lack of available datasets in this field. To address this, we developed a custom GPT-based talking bot to deliver cognitive tasks to participants. Our objective was to design a system that detects cognitive overload by monitoring subtle changes in human

movement. Our observations indicate that changes in human motion can be effectively detected using non-intrusive sensors. We achieved an accuracy of 68% for cognitive load classification using structural vibration signals with a Random Forest classifier, and 75.39% accuracy using acceleration signals. For person identification, we achieved an accuracy of 82% using structural vibration signals. Moving forward, we plan to conduct longer-term trials that also examine emotional factors and expand the participant pool to further validate and enhance system performance.

ACKNOWLEDGMENT

We would like to express our gratitude to the Department of Electrical Engineering (EE), SeNSE, and BSTTM at IIT Delhi for their invaluable support. We also extend our sincere thanks to Mr. Shivang Gaur for his assistance in data collection, as well as to all the participants of the study. This research is supported by the PMRF and the IEEE SPS Scholarship Grant.

REFERENCES

1. Sweller, J. 1988. "Cognitive load during problem solving: Effects on learning," *Cognitive science*, 12(2):257–285.
2. Chen, F., J. Zhou, Y. Wang, K. Yu, S. Z. Arshad, A. Khawaji, and D. Conway. 2016. *Robust multimodal cognitive load measurement*, Springer.
3. Korbach, A., R. Brünken, and B. Park. 2018. "Differentiating different types of cognitive load: A comparison of different measures," *Educational Psychology Review*, 30(2):503–529.
4. Chakraborty, M. 2022, "Cognitive Load Estimation," .
5. Ahmad, M. I., I. Keller, D. A. Robb, and K. S. Lohan. 2023. "A framework to estimate cognitive load using physiological data," *Personal and Ubiquitous Computing*:1–15.
6. Kyriaki, K., D. Koukopoulos, and C. A. Fidas. 2024. "A comprehensive survey of EEG preprocessing methods for cognitive load assessment," *IEEE Access*, 12:23466–23489.
7. Boffet, A., L. M. Arsac, V. Ibanez, F. Sauvet, and V. Deschodt-Arsac. 2025. "Detection of Cognitive Load Modulation by EDA and HRV," *Sensors*, 25(8):2343.
8. Martin, E. and R. Bajcsy. 2011. "Analysis of the effect of cognitive load on gait with off-the-shelf accelerometers," *Proc. COGNITIVE*:1–6.
9. Hart, S. G. 2006. "NASA-task load index (NASA-TLX); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, Sage publications Sage CA: Los Angeles, CA, vol. 50, pp. 904–908.
10. Abbad-Andaloussi, A., T. Sorg, and B. Weber. 2022. "Estimating developers' cognitive load at a fine-grained level using eye-tracking measures," in *Proceedings of the 30th IEEE/ACM international conference on program comprehension*, pp. 111–121.
11. Palinko, O., A. L. Kun, A. Shyrovkov, and P. Heeman. 2010. "Estimating cognitive load using remote eye tracking in a driving simulator," in *Proceedings of the 2010 symposium on eye-tracking research & applications*, pp. 141–144.
12. Khawaja, M. A., F. Chen, and N. Marcus. 2012. "Analysis of collaborative communication for linguistic cues of cognitive load," *Human factors*, 54(4):518–529.

13. Dasgupta, P., J. VanSwearingen, and E. Sejdic. 2018. ““You can tell by the way I use my walk.” Predicting the presence of cognitive load with gait measurements,” *Biomedical engineering online*, 17:1–18.
14. Ho, S., A. Mohtadi, K. Daud, U. Leonards, and T. C. Handy. 2019. “Using smartphone accelerometry to assess the relationship between cognitive load and gait dynamics during outdoor walking,” *Scientific reports*, 9(1):3119.
15. Nguyen, A., Y. Rothacher, E. Efthymiou, B. Lenggenhager, P. Brugger, L. Imbach, and A. Kunz. 2020. “Effect of cognitive load on curvature redirected walking thresholds,” in *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–5.
16. Pan, S., N. Wang, Y. Qian, I. Velibeyoglu, H. Y. Noh, and P. Zhang. 2015. “Indoor person identification through footstep induced structural vibration,” in *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, pp. 81–86.
17. Chakraborty, M. and S. Kar. 2023. “Enhancing Person Identification Through Data Augmentation of Footstep-Based Seismic Signals,” *IEEE Signal Processing Letters*, 30:1642–1646, ISSN 1558-2361, doi:10.1109/LSP.2023.3327650.
18. WU, Y., Y. DONG, S. VAID, G. M. HARARI, and H. Y. NOH. 2023. “Emotion recognition using footstep-induced floor vibration signals,” *STRUCTURAL HEALTH MONITORING 2023*.
19. Dong, Y., S. E. Kim, K. Schadl, P. Huang, W. Ding, J. Rose, and H. Y. Noh. 2024. “In-home gait abnormality detection through footstep-induced floor vibration sensing and person-invariant contrastive learning,” *IEEE Journal of Biomedical and Health Informatics*.
20. Chakraborty, M., B. Mukhopadhyay, S. Anchal, S. Kar, et al. 2025. “VibeGait: Enhancing Structural-Vibration based Gait Recognition using Vision,” in *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 1–5.
21. OpenAI. 2024, “ChatGPT (April 2024 version),” .
22. Inc., T. M. 2022, “MATLAB version: 9.13.0 (R2022b),” .
23. He, K., X. Zhang, S. Ren, and J. Sun. 2016. “Deep Residual Learning for Image Recognition,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 770–778.