ERUDITE: Human-in-the-Loop IoT for an Adaptive Personalized Learning System

Mojtaba Taherisadr⬤, Mohammad Abdullah Al Faruque⬤, *Senior Member, IEEE*, and Salma Elmalaki

Abstract—Thanks to the rapid growth in wearable technologies and advancements in machine learning, monitoring complex human contexts becomes feasible, paving the way to develop human-in-the-loop IoT systems that naturally evolve to adapt to the human and environment state autonomously. Nevertheless, a central challenge in designing many of these IoT systems arises from the requirement to infer the human mental state, such as intention, stress, cognition load, or learning ability. While different human contexts can be inferred from the fusion of different sensor modalities that can correlate to a particular mental state, the human brain provides a richer sensor modality that gives us more insights into the required human context. This article proposes ERUDITE, a human-in-the-loop IoT system for the learning environment that exploits recent wearable neurotechnology to decode brain signals. Through insights from concept learning theory, ERUDITE can infer the human state of learning and understand when human learning increases or declines. By quantifying human learning as an input sensory signal, ERUDITE can provide adequate personalized feedback to humans in a learning environment to enhance their learning experience. ERUDITE is evaluated across 15 participants and showed that by using the brain signals as a sensor modality to infer the human learning state and providing personalized adaptation to the learning environment, the participants' learning performance increased on average by **26%**. Furthermore, to evaluate ERUDITE practicality and scalability, we showed that ERUDITE can be deployed on an edge-based prototype consuming **75-mW** power on average with **100 MB** memory footprint.

Index Terms—Augmented reality (AR), concept learning, electroencephalography (EEG), Internet of Things, *Q*-learning, reinforcement learning, rule-based learning, virtual reality (VR), Wisconsin card sorting.

I. INTRODUCTION

THE DISRUPTION of face-to-face exchanges and the urgent digital transformation of our everyday activities, abide the COVID-19 hit have spurred us to shift many working paradigms to incorporate the new reality of the online and remote collaborative environment. In particular, many sectors, including education, workforce training, and the healthcare systems, had to integrate new technology to continue their services even with remote interactions [1], [2]. Nevertheless, this forced digital transformation of many sectors is expected to stay even in the post-COVID-19 era. In particular, in the education and workforce training sector, the e-learning market worldwide is forecast to surpass 243 billion U.S. dollars [3]. This paradigm shift needs a new approach to rethink the future learning environment to fit this new reality for remote and online learning environments without sacrificing the efficiency of the learning outcomes [4]. Moreover, even with regular in-person interaction, everyday life and complex work and learning environment make it difficult for humans to show extended focus, engagement, or attention on a specific learning task, directly affecting their performance in everyday tasks [5]. Hence, we need to rethink the future of the learning environment to incorporate the human learning state (LS) (focus, engagement, attention) as an integral part of its design.

Measuring and evaluating the learning outcomes is a cornerstone in the design of any learning environment. For example, traditional measures can take different forms in an educational environment, such as quizzes, exam scores, and teacher evaluation [6]. Different approaches have been proposed in the workforce environment, such as the learning-transfer evaluation model (LTEM), to measure if employees learned and whether they can or will perform the desired behavior on the job [7].

With the recent advances in IoT, machine learning models, and context-aware computing [8], we envision that the future of the learning environment has to be redesigned to incorporate a *real-time* measurement of the human LS with *real-time* assessment and provide adequate feedback to the human and the instructor (whether remote or online or in-person learning) to improve the learning outcomes [9], [10]. Accordingly, in this article, we ask the following questions.

Q1: Using IoT technologies (wearable sensors and edge-cloud computing), can we understand the learning process and infer the human LS at run-time?

Q2: Can we integrate this LS signal as a sensor modality in the IoT-based learning environment to make a real-time adaptation and improve the LS of a human?

Q3: Can we measure the effectiveness of these adaptations and tune them to provide a more personalized learning experience?

These questions have multifold challenges. A fundamental challenge is understanding how the human brain learns a new concept. Indeed, recent advances in neuroscience have opened the gate to unveil fundamental processes in the human brain, such as the ability to generate emotions, memories,

and actions [11]. In particular, electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) were the two primary approaches used to record and measure the human brain state under learning processes [12], [13]. While these approaches provide necessary insights into the human brain, the main drawbacks are that they must be conducted in a laboratory setting, are computationally costly, and are inadequate in a natural learning environment, such as education or workforce training, where IoT-based systems are more adequate. Hence, other approaches used real-time signals from physiological sensors available in commodity wearables to correlate learning outcomes with different physiological states and hence can use machine learning models to predict the learning experience [14], [15], [16]. However, these approaches measure the learning indirectly as a correlation to other physiological states, such as engagement, attention, stress, focus, or comparatively, under different stimuli. Moreover, these approaches rarely close the loop by providing real-time adaptation to improve human learning.

Nevertheless, another challenge to answering the aforementioned questions is that even if we can quantify the human LS outside the laboratory settings using IoT devices (such as wearables) and independently from other physiological states, learning varies from time to time and from person to person. The state of learning or memory formation varies over time for any individual, partly due to moment-to-moment fluctuation of brain state. Moreover, varying levels of alertness, attention, and motivation likely contribute to fluctuating brain states for learning [13]. In addition, recent research in the human-in-the-loop IoT systems has shown that humans have intrinsic inter-and intra-variation. Hence, we cannot have one model-fits-all approach in designing real-time adaptations to improve the human learning [17], [18], [19], [20], [21].

In this article, we propose *ERUDITE*, an IoT-based system to quantify the human LS using a noninvasive wearable technology by decoding the brain signals under different learning approaches. We then use the human LS signal as another sensor modality in the IoT-based learning environment to design human-in-the-loop adaptations to enhance human learning.

Contribution: Our contributions in this article can be summarized as follows.
1) Exploring two concept-learning-based experiments to determine the human LS as a sensor modality (*Q1*).
2) Exploiting recent IoT technology (wearable EEG devices) to find the correlations between EEG signal dynamics and human LSs (*Q2*).
3) Designing ERUDITE, an IoT system that integrates the wearable EEG device as a sensor modality with an edge-based prototype to provide real-time adaptation for the learning environment and online feedback to the human to improve the learning (*Q3*).

II. BACKGROUND AND RELATED WORK

In this section, we briefly discuss some background related to the philosophy of learning that we will use as a basis to design *ERUDITE*. Afterward, we list recent research approaches in quantifying learning using brain signals.

A. Concept Learning Theory

Concept learning is defined by Bruner [22] as "the search for and listing of attributes that can be used to distinguish exemplars from nonexemplars of various categories." Thus, concept learning is a strategy that requires a learner to compare and contrast groups or categories that contain concept-relevant features with groups or categories that do not contain concept-relevant features. Concept learning must be distinguished from "learning by reciting" something from memory (recall) or "discriminating between two things" by refining concepts (discrimination). In this article, we borrow from philosophy and psychology literature to exploit two types of concept learning, namely, 1) rule-based concept learning and 2) explanation-based concept learning. We briefly explain them below.

1) *Rule-Based Concept Learning* has initiated with cognitive psychology and early learning models that could be implemented with a computer language using conditional statements, such as if: then, which produces rules. Computer models utilize training data and a rule-based approach as input, which results from a rule-based learner to produce a more accurate data model. A rule-based system can be used in concept learning when the stimuli are confusable instead of simple. When rules are used in learning, decisions are made based on properties alone and rely on simple criteria that do not require a lot of memory [23]. For instance, a radiologist uses the rule of differences in brightness to categorize the X-ray images. Using the rule, the radiologist decides which X-ray image belongs to what category.

2) *Explanation-Based Concept Learning* suggests that a new concept is acquired by experiencing examples of it and forming a basic outline. In particular, by observing or receiving the qualities of a thing, the mind forms a concept that possesses those qualities [24]. The revised explanation-based learning model integrates four mental processes – generalization, chunking, operationalization, and analogy. *Generalization* is the process by which the characteristics of a concept that are fundamental to it are recognized and labeled. For example, birds have feathers and wings. Hence, anything with feathers and wings will be identified as a "bird." When information is grouped mentally, whether, by similarity or relatedness, the group is called a *chunking*. An example of chunking occurs in phone numbers; a phone number sequence of $4-7-1-1-3-2-4$ would be chunked into $471-1324$ to remember them more manageably.

A concept is *operationalized* when the mind can turn abstract concepts into measurable observations. Operationalization is a process of defining the measurement of a phenomenon that is not directly measurable, though its existence is inferred by other phenomena. For example, in medicine, the phenomenon of health might be operationalized by one or more indicators like body mass index or tobacco smoking. *Analogy* is the recognition of similarities among potential examples. For example, an atom is like a miniature solar system and a database is like a filing cabinet.

In this article, we will borrow from the rule-based concept learning theory due to its deterministic approach in identifying if a rule is learned or not to get insights into how to measure the human LS. Afterward, we will use these insights to analyze the explanation-based concept learning since its concept is more applied in learning environments.

B. Learning Quantification

Two bodies of work use brain signals to quantify learning. One approach uses fMRI and the other is EEG signals. We summarize them below.

1) fMRI: Yoo et al. [13] designed two experiments to measure the LS. One experiment examined whether such brain states that were good or bad for learning concepts could be identified. In another experiment, they examined whether such brain states, detected by real-time fMRI on a moment-moment basis, could be used to trigger concept presentation with the hypothesis that concepts triggered by good brain states would be better remembered than concepts triggered by bad brain states. Such a finding would offer evidence of the ability to monitor online whether a person is optimally prepared to learn and the ability to use fMRI to causally enhance human learning in the sense that the real-time fMRI-measured brain state caused the concept presentation. Toni et al. [25] exploited fMRI to measure changes in blood oxygen level-dependent (BOLD) signals throughout learning. Their analysis found linear and nonlinear changes of BOLD signal over time in prefrontal, premotor, and parietal cortex areas.

In particular, Babini et al. [26], using an EEG device, designed an experiment to analyze the effect of virtual reality (VR) conditions on students' learning ability and physiological state. Their results showed that students' learning ability was increased in the 3-D condition compared to that in the 2-D condition. In this study, they compared the engagement of the brain in different presentation modes to correlate it to the learning outcome. Xu and Zhong [27] reviewed portable EEG devices in educational research and envisioned that portable EEG would be employed extensively in the education field in the near future.

In this article, as we envision a human-in-the-loop IoT learning environment, it is crucial that we consider wearable devices that can be integrated with edge devices. Hence, we build upon recent work in the literature that uses EEG from wearable devices for learning measurement.

C. Human-in-the-Loop Approaches for Learning

Changing the learning environment to incorporate the human and improve the human learning experience has been addressed in the area of Smart Classrooms. Chamba-Eras and Aguilar [28] proposed the utilization of augmented reality (AR) and an agent that provides services of AR to display and design augmented scenarios in a smart classroom. In addition, they designed an experimental environment to evaluate the utilization and impact of AR in it. Their results showed that the students' criteria of motivation, learning curve, and memorization were substantially improved in the experimental group. Ahuja et al. [29] proposed "EduSense," a system
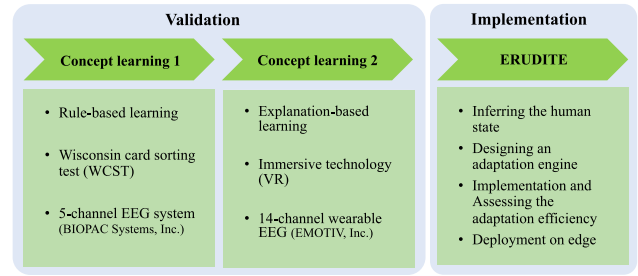

Fig. 1. Roadmap of ERUDITE validation and implementation. Two concept learning-based experiments are conducted for validation and then implementation is done based on the insights gained from those experiments.

that captures both audio and video streams of the instructor and students to analyze effective instruction, including hand raises, body pose, body accelerometry, and speech acts, and provides analytical feedback to the instructor. While the education sector has already taken technological leaps over the last decades with the introduction of smart classroom concept [29], [30], [31], efficient real-time measurement of the learning outcomes that incorporate the student's mental state has to be considered [9], [18].

This article is organized as follows; first, we explore two concept learning approaches, rule-based and explanation-based, in Sections III and IV, respectively, to get insights into how to measure the LS of the human. In Section V, we describe ERUDITE framework for a human-in-the-loop IoT system for a learning environment. We evaluated ERUDITE on 15 human subjects and discussed the deployment of ERUDITE on an edge-based prototype in Section VI. Finally, we conclude this article with some discussion and future work in Section VII. Fig. 1 illustrates the roadmap of ERUDITE, where in the validation phase we exploit two concept-learning experiments as discussed in Sections III and IV to gain insights into the correlations between learning and brain signals. Then we utilize these insights in the implementation phase to design ERUDITE as discussed in Section V.

III. RULE-BASED CONCEPT LEARNING

As we highlighted in the background in Section II, cognitive psychologists and philosophers have paid particular attention to concepts that identify kinds of things–those that classify or categorize objects—and such concepts that are named rule-based learning [23]. In this section, we utilize the rule-based learning concept to explore the ability to understand and quantify learning through EEG signals. In particular, we choose the Wisconsin card sorting test (WCST) to represent the rule-based learning [32].

A. Wisconsin Card Sorting Test

The WCST is a neuropsychological test frequently used to measure higher level cognitive processes, such as perseverance, attention, abstract thinking, and learning [32]. The WCST is a widely recognized and standardized neuropsychological assessment tool. It has a rich historical background, serving as a benchmark for comparing our research results with existing literature [32], [33]. The test's reliability and
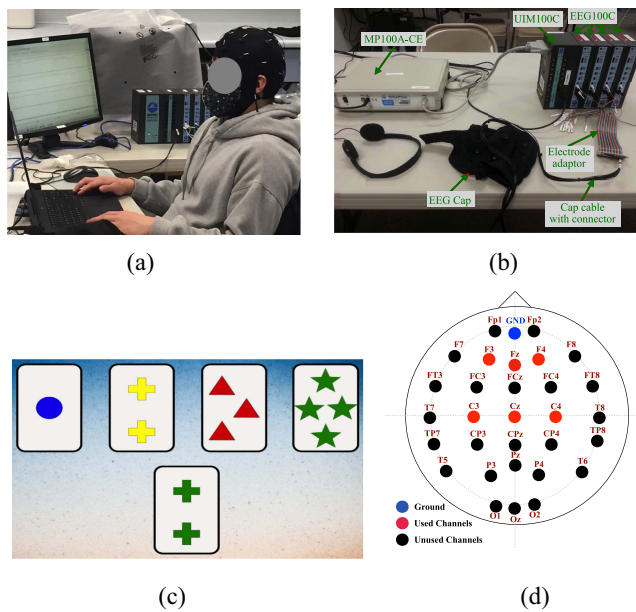
(a)   (b)

(c)   (d)

Fig. 2.   WCST experiment setup and EEG device. (a) EEG cap worn by a participant while performing the WCST. (b) EEG amplifier, data acquisition unit, and accessories. (c) Instance of the WCST environment. In this figure, the rule is color (green). (d) EEG channels' names and locations.

Fig. 3.   Total number of correct answers for each participant in WCST across two different sessions.

accessibility make it a solid choice for evaluating cognitive and executive functions. This availability of data supports the reproducibility of our findings and facilitates further research [34], [35]. Cognitive flexibility, the ability to switch between mental tasks, is a crucial aspect of human learning, and the WCST objectively assesses it, minimizing the influence of bias or subjective judgment [36]. The WCST consists of two card packs, each containing four stimulus cards and 128 response cards. The cards are of various geometric shapes in different colors and patterns. The participants are expected to accurately sort every response card with one of four stimulus cards through the feedback (right or wrong) given based on a predefined rule that is oblivious to the participant. Among various versions, the version of WCST with 128 cards developed by Heaton is what we adopted [37].

Fig. 2(c) depicts the computerized version of the WCST environment. The first row is the four stimulus cards that the player has to choose among them, and the second row is the test card that has to be matched with one of the stimulus cards based on some rule that the participant needs to learn. For example, if the rule is "Matching Color" and the participant chooses the pattern card, the feedback will be "wrong." Participant has five chances to find out (*learn*) the rule and choose the right card. If the participant selects the right card in any of the five moves, the move is called the "correct move," and one of the 128 card rounds ends. If, after five tries, the participant cannot find out the right rule, the current round ends without any right answer. Then the sorting rule changes discreetly from, i.e., color to form or pattern of figures. This rule change does not follow any pattern, and it happens randomly without the participants being informed. The number of the correct moves represents the times the participant could learn the correct rule out of 128 rounds of the game [32].
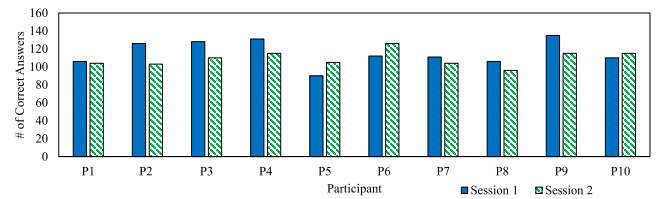
B. EEG Data Collection

As the purpose of this experiment is to get insights on how to quantify the learning event, we used a 5-channel research-grade EEG system (BIOPAC Systems, Inc. [38] shown in Fig. 2(b) with a sampling rate (Fs) of 200 Hz for data collection. Channel $Fpz - 1$ was selected as the ground for data acquisition setup, and the electrodes' impedance was clamped at less than 5 KΩ. Electrode locations were based on the 5-channel subset of 24-channel $10 - 20$ systems and were chosen to be over the frontal areas of the skull because it is the area in the brain that is responsible for cognitive processes, such as executive function, attention, memory, and language [39]. The electrodes are attached to Fz, $F3$, $F4$, $C3$, and $C4$ channels. Fig. 2(a) illustrates the setup attached to one of the participants. Fig. 2(b) shows the EEG device setup and associated accessories. Fig. 2(d) presents the EEG channels' locations. We used the *AcqKnowledge* recording platform to record the EEG signals from these channels [38].

We conducted the WCST experiment on ten healthy participants, including six males and four females, all between 20–30 years old. The average and standard deviation of the age of the participants were 27 and 2.36, respectively. Each participant was asked to do two separate sessions for a total of 20 sessions across all participants. Each session was done on different days and at different times of the day. Before the experiment, the color blindness test is used (e.g., by the Ishihara Plate Test [40]) to ensure it does not influence the test result.

C. Analysis and Correlation of EEG Signal to Learning

1) Memorization Effect: A central assumption in rule-based concept learning is that there is no effect of memory or experience in the learning process. Hence, no matter the number of times a participant conducts the WCST, the performance is independent of the previous sessions. Thus, in rule-based concept learning, the participant only follows the current rule of the game. Fig. 3 illustrates the number of correct moves each participant took in two separate sessions. Fig. 3 represents the number of correct moves for each participant in session 1 and session 2 of the experiment. The total correct move for all participants in session 1 and session 2 are 1154 and 1105, respectively. Overall, the correct moves change by ±4% between sessions. As Fig. 3 shows, from session 1 to session 2, the number of correct moves does not follow a particular pattern for participants—decrease for some participants and increase for others. This indicates that
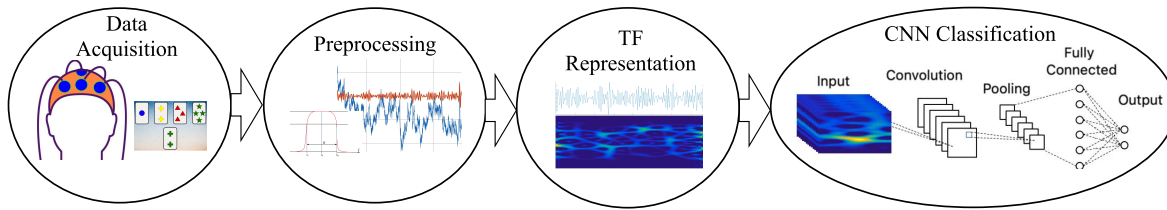
Fig. 4.   Data acquisition and analysis of the rule-based concept learning experiment. After acquiring the data and preprocessing, TF representation is applied to each window of the EEG signal. Then TF images with labels acquired from wrong and correct moves of the WCST experiment are fed to the CNN model for classification.
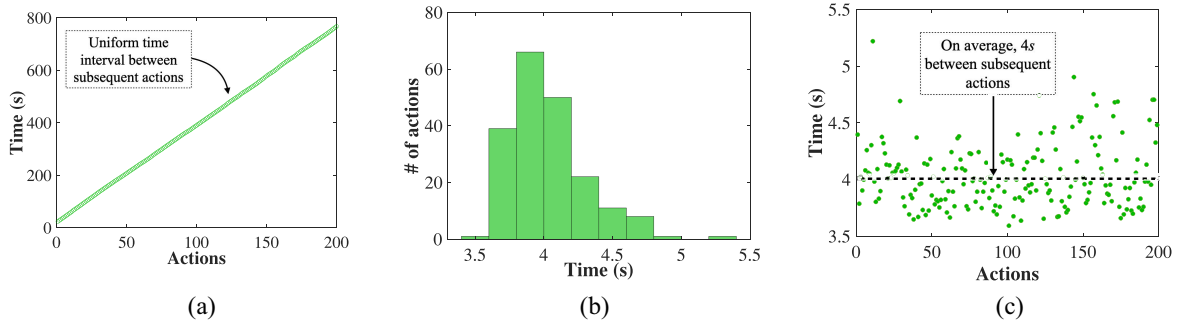

Fig. 5.   Windowing size analysis averaged across all participants to segment the EEG data for WCST. (a) Timestamp for the actions taken averaged across all participants during the WCSG experiment. The linear growth of the timing depicts the uniformity of the time taken for subsequent actions. (b) Histogram of the average time the participants took to take one action in WCST. The horizontal axis is the time intervals in seconds, and the vertical axis is the number of actions taken in each time interval. (c) Time taken per action averaged across all participants in WCST.

repeating the WCST has no effect on the result (as experience or memorization).

2) Effect of Correct Moves on EEG Dynamics: We analyzed the collected EEG data to evaluate the correlation between the EEG dynamics and the rule-based concept learning experiment. In particular, how to use the EEG signal to infer whether the participant learned the rule or not. An overall pictorial figure of our analysis approaches is shown in Fig. 4. Below we describe the details of this analysis to detect this correlation.

1) *Filtering and Facial Artifacts:* EEG data is preprocessed using a high-pass filter (HPF) with a cut-off frequency of 0.5 Hz for removing trends and 40 Hz to remove the high-frequency components. To remove facial artifacts, such as jaw movement, eye blink, and eye movement, we applied independent component analysis (ICA). ICA is a computational method for separating a multivariate signal into additive subcomponents. The assumption is that the subcomponents are statistically independent of each other, which applies to EEG signals and facial artifacts [41], [42].

2) *Windowing:* Since the WCST has a very short response time, we need to capture this response event carefully. Hence, we applied a 4 s window with 1 s overlap. We segmented the EEG data into multiples of 4s data using this 4 s window. The reason is that, on average, every participant takes one action every 4 s. Fig. 5(a) depicts the timestamp of the taken actions averaged across all participants. As Fig. 5(a) shows, the timing grows linearly, which refers to the uniformity of the time a participant takes between subsequent actions. We use this outcome as the base for segmenting the EEG

signal. With this window size, aligning or warping the EEG data to fit the segments is not required. Due to the uniformity of the time taken between subsequent actions, we chose a fixed window size. In particular, Fig. 5(b) shows the distribution of the time intervals between subsequent actions averaged across all participants where we observe more than 80% of the actions are taken between 375 and 450 ms, which, in turn, illustrates that window size 4 s is the dominant window size for the participants. Fig. 5(c) shows the scatter-plot of the time all participants take between subsequent actions.

3) *Spectro-Temporal Representation of the EEG Signal:* After preprocessing, we utilize time-frequency (TF) representation to analyze the collected EEG signals in the temporal and spectral domains. TF could be considered a nonstationary signal analysis with frequency content varying with time. TF is a suitable representation for nonstationary and multicomponent signals, such as EEG, which can simultaneously describe the given signal's energy distribution over time and frequency space. TF selection for EEG signal representation is the first step in designing an appropriate representation. A proper TF highlights the nonstationarities in the input signal that enables the system to discriminate between variations of the signal in both temporal and spectral domains. We use the TF representation to investigate the spectro-temporal correlations between the learning event and EEG signals. In particular, we used the smoothed Wigner-Ville distribution (smoothed-WVD) to create the TF representation since, in practice, it has been shown to improve the quality of representation due to its ability to reduce the cross-term interference for a signal with multiple
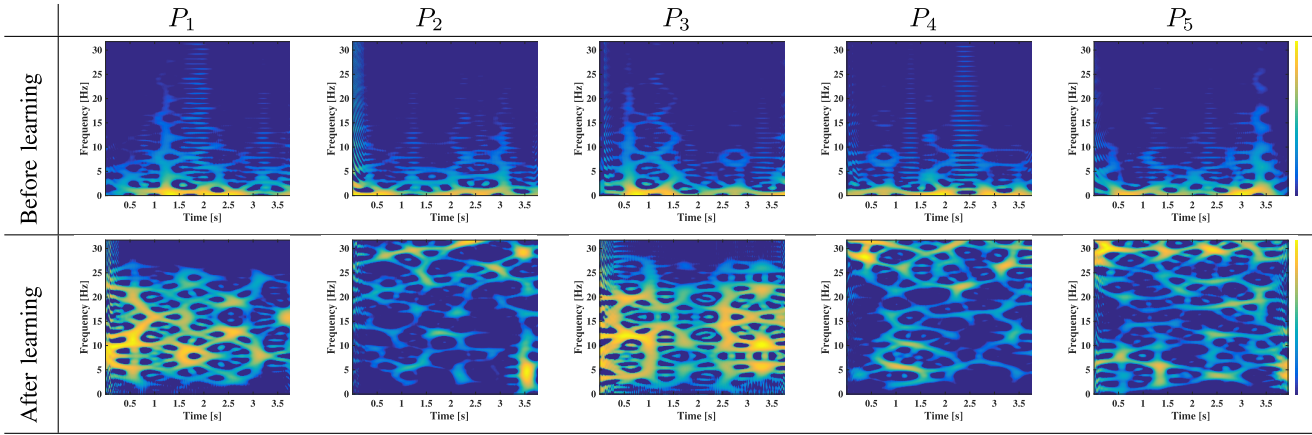
Fig. 6. TF representation using smoothed-WVD of EEG signal from Fz channel. Brighter color refers to higher energy in the associated frequency in Hz and time in seconds, and darker colors mean lower energy for the associated frequency in Hz in the given time in seconds. (a) First row represents the TF representation of the EEG signal for five participants before they learn the rule in WCST. (b) Second row represents the TF representation of the EEG signal for the same five participants after they learn the rule in WCST by taking five correct actions.

frequency components [43].[1] Fig. 6 illustrates the TF representation of the EEG signal from the Fz channel for five participants.[2] In particular, the first row in Fig. 6 depicts 4s duration TF representation of the EEG signal before the participants learn the rule. The second row in Fig. 6 shows the same duration of the EEG signal after participants find out (*learns*) the rule and can take five consecutive correct actions, which is considered the duration they have completely mastered the rule. After participants learn the rule and take consecutive correct moves, we observe that the representation contains more activities in the higher spectral component. To quantify this observation, we measured the spectral power of the high-frequency sub-bands (10–25 Hz) of the EEG signal from Fz channel before and after the participant learns the rule of the WCST as shown in Table I.[3] Results in Table I show spectral power increases on average by 76.52% after participants learn the rule of the WCST.

D. Classification of Learning Events in WCST

As we mentioned earlier, we aim to use a representation of the LS as a sensor modality in a learning IoT system. One approach to infer this LS is to use a supervised learning model. In particular, we want to infer two classes; learning and not-learning events; where the "learning" label is assigned if there are five consecutive correct moves in the WCST, and the not-learning label is assigned if there is one wrong move. We applied three classifiers with different computation complexity to do this classification.

[1]Methods for TF representation can be categorized in six groups as follows: 1) Gaussian kernel; 2) Wigner–Ville distribution (WVD); 3) spectrogram; 4) modified-B; 5) smoothed-WVD; and 6) separable kernel. Reduced interference approaches, such as smoothed-WVD, can improve the representation quality by reducing the effect of the cross-terms [44].

[2]We did the same analysis for all participants and depicted five participants in the Fig. 6 for visualization purposes.

[3]We use v^2Hz^{-1} unit for spectral power, as the EEG signal is defined in terms only of a voltage and there is no unique power associated with the stated amplitude.

TABLE I
SPECTRAL POWER OF THE HIGH-FREQUENCY SUB-BANDS (10–25 HZ) OF THE EEG SIGNAL FROM THE Fz CHANNEL BEFORE AND AFTER THE PARTICIPANT LEARNS THE RULE OF THE WCST

Participant	Before learning		After learning	
	Session 1 (v^2Hz^{-1})	Session 2 (v^2Hz^{-1})	Session 1 (v^2Hz^{-1})	Session 2 (v^2Hz^{-1})
Participant 1	0.094	0.087	0.223	0.196
Participant 2	0.079	0.083	0.193	0.213
Participant 3	0.138	0.112	0.251	0.217
Participant 4	0.201	0.189	0.297	0.257
Participant 5	0.147	0.183	0.261	0.239
Participant 6	0.129	0.198	0.272	0.298
Participant 7	0.085	0.097	0.179	0.194
Participant 8	0.139	0.172	0.199	0.216
Participant 9	0.091	0.118	0.183	0.213
Participant 10	0.169	0.196	0.214	0.268
Average	0.127	0.144	0.223	0.231

1) *Support Vector Machine (SVM) Model:* We designed an SVM discriminative linear classifier with a Gaussian kernel using the spectral and temporal features from the preprocessed EEG data. Features used to train the SVM model are a) autoregressive (5 features), which represents an estimation of temporal characteristics of the EEG signal; b) four features extracted from each sub-bands of the Wavelet decomposition of the EEG signal; and c) short-time Fourier transform (STFT) decomposition, including four features. We applied five levels of "Haar" Wavelet decomposition to each EEG window (which provides five sub-bands) and extracted four features from each sub-band. In total 29 features were extracted.

2) *Convolutional Neural Network (CNN):* We designed a supervised CNN model to classify the 4s windows of the TF representation as an input image (as shown in Fig. 6) into two classes; learning and not-learning events. We used four main building blocks in the CNN model: a) convolution; b) pooling; c) rectified linear unit (ReLU); and d) a fully connected layer. The designed model has 4 convolutions, 3 max-pooling, 1 flatten, and 1 fully

TABLE II

ACCURACY (%) OF THE SVM, CNN, AND DSTCLN CLASSIFIERS FOR TWO SEPARATE SESSIONS OF WCST AVERAGED ACROSS TEN PARTICIPANTS. THE SENSITIVITY OF THE CLASSIFIERS AND VARIANCE OF THE 10-CV RESULTS ARE AVERAGED ACROSS SESSIONS 1 AND 2

Channel	SVM				CNN				(DSTCLN)			
	Session1	Session2	Sensitivity	Variance	Session1	Session2	Sensitivity	Variance	Session1	Session2	Sensitivity	Variance
Fz	58.36%	51.53%	59.35%	8.17%	81.03%	77.16%	84.27%	6.26%	84.81%	83.35%	86.18%	5.82%
$F3$	67.42%	70.12%	73.64%	11.26%	71.92%	72.73%	75.23%	7.56%	73.23%	75.32%	78.19%	8.12%
$F4$	59.13%	65.89%	64.83%	9.27%	78.73%	73.45%	80.34%	8.81%	79.87%	77.16%	80.42%	9.14%
$C3$	49.83%	57.61%	45.72%	13.39%	69.81%	65.32%	66.38%	11.72%	72.63%	70.31%	74.13%	10.21%
$C4$	56.27%	62.18%	51.67%	16.31%	72.31%	74.28%	75.13%	8.21%	75.18%	77.07%	78.68%	7.62%

connected layer as described in [45]. The inputs fed to the CNN model are the images of the TF representation of the $4s$ window with the label learning or not-learning as explained above.

3) *Long Short-Term Memory (LSTM)-Based Model:* We exploited a recent LSTM-based model in the literature, which is a deep spatio-temporal convolutional bidirectional LSTM network (DSTCLN) [46] to classify the EEG signal into two classes; learning and not-learning. Details of the DSTCLN model are provided in Appendix A.

In all these three classifiers, tenfold cross-validation (10-CV) was applied.[4] We report the results of the classification accuracy of the three classifiers in Table II. In particular, the three classifiers were used on the EEG data collected from the five channels that we considered from the frontal lobe. In Table II, we report the average across all the participant populations in the two different sessions, as we explained before. As Table II illustrates, the classification accuracy of DSTCLN outperforms the classification accuracy of SVM per EEG channel. Moreover, in DSTCLN, the accuracy of classification using the Fz channel is higher than the other channels. Table II also illustrates the sensitivity and the variance of the 10-CV results of the classifiers. As the classification results of Table II show, channel Fz outperforms other channels and depicts the correlation between the inferred learning class (learning or not-learning) and that region of the brain.

In this analysis, we studied the EEG channels separately and not a combination for two reasons.

1) We need to localize the part of the brain related to learning new concepts. For this purpose, we considered each channel's performance independently to determine the regional correlation between EEG channels' location and firing neurons activated during the learning event.

2) We need to reduce the computation cost of processing the EEG signal as we aim to use it in IoT systems. Hence, we find the best performing channel instead of considering all channels or a different combination.

Moreover, although the DSTCLN model performs better than the CNN model, it comes with a higher computation cost. In

---

[4]10-CV divides the total input data of n samples into ten equal parts. In every iteration, 1 part is considered a test sample set, and the remaining nine parts are considered for validation and training sample sets. There is no overlap between the test sample set (10% of data) with the validation and training sample set (90% of data).

particular, DSTCLN has ≈ 30 million parameters compared to ≈ 28 million parameters in the CNN model. More details on the timing analysis and the processing overhead of the EEG data will be discussed in Section VI.

E. Insights From Rule-Based Concept Learning

From the rule-based concept learning experiment, we gained three main insights.

1) Using TF representation, we observed that consecutive correct moves correlate with higher spectral activities in the higher frequency ranges, as shown in Fig. 6. Moreover, the spectral power of the EEG signal from the Fz channel shows an increase of 76.52% when the learning event occurs (five consecutive moves) as reported in Table I (addressing *Q1* in the introduction).

2) Results from three different classifiers (DSTCLN, CNN, and SVM) across two different sessions of WCST averaged on ten participants showed that the Fz channel can be used to classify the learning versus not learning events (addressing *Q1* in the introduction).

3) The total correct moves by each participant in two sessions shows no experience or memory involved in the results. As illustrated in Fig. 3, there is on average $\pm 4\%$ change between sessions in the total number of the correct moves for all participants.

We will use the insight we gained from the rule-based learning experiment that learning event correlates positively with the high-frequency sub-bands of the frontal lobe channels of the EEG signal to study the explanation-based concept learning as explained in Section IV.

IV. EXPLANATION-BASED CONCEPT LEARNING

While the rule-based concept learning experiment gave us the necessary insights to classify the learning event, there are two pivotal points that we need to consider.

1) Humans can learn through techniques other than rule-based concept learning.

2) Although we needed to use a high-accuracy EEG cap [Fig. 2(b)] to exploit the feasibility of classifying the learning events, we need to consider more socially acceptable wearable devices as we envision human-in-the-loop IoT learning system [18].

As highlighted in the background section (Section II), we exploit another concept learning approach called "explanation-based concept learning." Explanation-based learning suggests that humans learn new concepts by experiencing examples and forming a basic outline. Accordingly, explanation-based learning can be viewed as presenting concepts in some form or "lecturing" new materials to humans. Hence, in this section, we pick teaching materials explained through a video and present them to different participants to understand the effect of explanation-based concept learning on brain activity in different presentation modalities.

A. Presentation Modalities for Learning

By the definition of this learning methodology, no specific event indicates the learning event akin to choosing the correct rule in rule-based concept learning. Hence, we need to assess whether the human retained and learned the content of the videos through some questionnaires while monitoring the EEG signals under different presentation modalities. In particular, we used two modalities of the same teaching material: 1) traditional modality and 2) immersive modality. The traditional modality version of videos is recorded under the traditional 2-D videos. The immersive modality version of the videos is recorded as 3-D videos. The 3-D video content is created by converting video from 2-D to 3-D using tools, such as the Ani3D [47], which creates imagery for each eye from one 2-D image. This 2-D-to-3-D conversion adds the binocular disparity depth cue to digital images perceived by the brain, thus, significantly improving the immersive effect when broadcasted in VR.

While previous research has explored the effect of 3-D and, more specifically, the impact of VR on brain engagement [26], what we aim to do through this experiment is to be able to quantify the learning event in the explanation-based concept learning using a wearable device in the two different presentation modalities (traditional and immersive).

B. Data Collection Using Wearable Device

We recruited 15 healthy participants for this experiment. Since caffeine affects brain activity and causes inconsistent results, participants were asked not to consume beverages containing caffeine within 24 h before the experiments [48].

We chose different topics depending on the participant's backgrounds to ensure they did not have prior knowledge of the presented material in the video. In particular, we picked videos covering topics on biology [49], modern architecture [50], and space [51]. The duration of these videos is $\approx$ 7.5 min, $\approx$ 7 min, and $\approx$ 12 min, respectively. We divided each video into two parts and converted them into 3-D, meaning that we have two versions of the same video content in 2-D and 3-D. Hence, we have six videos (2-D and 3-D) and six sets of a questionnaire for each video.

We recorded the EEG signal using EMOTIV EPOC_X [52], a commodity EEG wearable device. We used the 14 Channel EEG headset with a sampling frequency of 256 Hz. Fig. 7(a) shows the wearable EMOTIV EPOC_X device placed on the participant's head with the Oculus device. Electrode

TABLE III
PARAMETERS OF EMOTIV EPOC_X EEG SETUP FOR EXPLANATION-BASED CONCEPT LEARNING

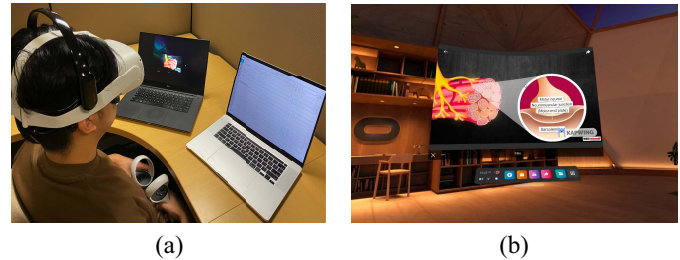| Parameter	Description
Amplifier	14-channel device (EMOTIV [52])
Sampling frequency	256 Hz
Bandwidth	0.16 − 43 Hz
Electrode arrangement	14-channel subset of 24-channel 10−20 systems (Figure 7a)
Ground reference	P3, P4
Electrode type	Ag/AgCl

Fig. 7. EMOTIV Epoc_x and Oculus devices worn by a participant for VR learning environment. (a) EMOTIV Epoc_x and Oculus devices are worn by a participant. The left screen in the figure casts the view of the participant on the Oculus device and the right screen presents the EEG signal collected from the EMOTIV device. (b) Screenshot of the view of the participant while watching biology content on an Oculus device with an office background.

arrangement was a 14-channel subset of 24-channel 10−20 systems where the P3 and P4 channels are the ground references. Table III provides the details and settings of the EMOTIV EPOC_X EEG device. For the 3-D videos, we used the Oculus VR setup and streamed the 3-D video on it. Fig. 7(b) shows an example of what a participant visualizes in the VR environment while watching one of the 3-D videos.

Before the experiment, we introduced the 3-D video presentation broadcasted on the VR device (Oculus) to participants to familiarize them with the technology to reduce any possible effects of facing new technology in the EEG signal. The experiment protocol started by asking the participants to watch the traditional 2-D video or the immersive VR 3-D video. Afterward, the participants were asked to answer three questions related to the content. After a 1-min rest period, the participants switched to the other modality (2-D to VR or VR to 2-D) of the same content. Three questions were then asked again, followed by another rest period for 1 min. The order of presenting the 2-D and VR videos is random. There is no emphasis on delivering one modality first, meaning that we show the 2-D video first and then the VR video; in other cases, it was vice versa. We continued this protocol to collect EEG signals from the participants (along with the responses to content-related questions) with the rest of the 6 sets of videos, providing the participants with 1 min of rest between watching the videos. To ensure the repeatability of the result, we made two sessions separated by one month for the 15 participants and changed the questionnaire in the second session.

TABLE IV
Fractal Dimension of EEG Signals With Two Presentation
Modalities (Traditional (2-D) and Immersive (VR)) Averaged
on All 15 Participants. The Lecture Contents Are
Biology, Architecture, and Space

Content and Presentation Modality	Fractal dimension	
	Session 1	Session 2
Biology (2D)	1.7319	1.7311
Biology (VR)	1.7393	1.7439
Architecture (2D)	1.7302	1.7312
Architecture (VR)	1.7405	1.7398
Space (2D)	1.7339	1.7311
Space (VR)	1.7421	1.7409

C. Analysis

Unlike rule-based concept learning, no correct/wrong moves during the experiment can be used as classification labels in explanation-based concept learning. Hence, we need another approach to quantify the learning event in the EEG data. One approach is to compare the complexity of brain activity during the different types of videos and their correlation with the performance in the questionnaire.

1) Brain Engagement Using Fractal Dimension: One method that has been implemented to explore and compare the complexity of the EEG signal during the learning experiments is fractal dimension [53].[5] Various methods have been developed to calculate the fractal dimension, mainly based on the entropy concept. In this experiment, we used the box-counting method to calculate the fractal dimension [53] on the recorded time-series EEG data. In particular, we used the $F3$ and $F4$ channels and averaged their values for our analysis.[6]

After preprocessing (filtering and noise removal, including HPF and ICA, as explained in Section III-C, we applied the box-counting algorithm and examined the fractal dimension of all the 15 participants across both sessions using the two presentation modalities. In this step, we do not window the EEG signal because there is no exact label for learning events. Thus, we process the EEG signal as a whole. In Table IV, we report the average fractal dimension across all 15 participants, and Fig. 8 shows the fractal dimension for the 15 participants across the two sessions. In general, we observe that the fractal dimension of the EEG signal is higher in the immersive VR modality compared to the traditional 2-D videos. Since the fractal dimension reflects the complexity of the signal, this result indicates that the EEG signal is more complex in response to VR visual stimuli than traditional 2-D visual stimuli. In other words, the human brain becomes more engaged with a stimulus when presented in the immersive VR compared to the traditional 2-D.

While the fractal dimension analysis gives us insights about brain engagement as a *comparative* approach between the

[5]Fractal dimension, as the primary quantitative measure of fractal theory, indicates the complexity of the process in which greater values of a fractal dimension reflect greater complexity of the object.

[6]EMOTIV device does not have an Fz channel, so we relied on the neighbor channels $F3$ and $F4$ and averaged their output signal to have an approximation of the region.

traditional 2-D and the immersive VR during the learning, we continue the analysis by investigating whether the same insights from the rule-based concept learning can be applied here.

2) Correlation of Spectral Power of the EEG Signal and the Learning Performance: We use the insight from the rule-based learning experiment that the learning event correlates positively with the high-frequency sub-bands of the frontal lobe channels of the EEG signal. In particular, we want to investigate whether the same insight can be applied to explanation-based concept learning. To this end, we measure the power of the spectral component of the EEG signal (10–25 Hz) band of the channels $F3$ and $F4$) in the two presentation modalities and its correlation with the participants' answers in the questionnaire as our ground truth. Table V provides the spectral power of the high-frequency sub-bands of the EEG signal in the two sessions of the experiment. As Table V shows, there is an increase in the spectral power from the traditional 2-D presentation mode to the immersive VR mode in different video contents and sessions. The increase in the spectral power of the high-frequency sub-band is aligned with the increase of the performance ($\approx 20\%$) provided by the questionnaire at the end of each content.

D. Insights From Explanation-Based Concept Learning

By conducting this experiment, we aimed to investigate the correlations between the EEG signal and explanation-based concept learning. The questionnaire served as the ground truth and showed that participants performed better during the immersive VR presentation (by 20%), which correlates with an increase in fractal dimensions and spectral power of the higher spectral components of the EEG signals from the frontal lobe. This insight aligns with the insights from rule-based learning.

The remainder of this article will discuss how we utilize the insights gained from rule-based and explanation-based experiments about learning to design ERUDITE, a human-in-the-loop IoT learning system.

V. ERUDITE (Human-in-the-Loop IoT Learning System)

In the previous sections, we gained some insights into how the human brain can react to learning events in two different learning approaches, rule-based concept learning, and explanation-based concept learning. We will utilize these collected insights on cues in learning, where the correlation occurs between EEG dynamics and learning to design ERUDITE, a human-in-the-loop IoT learning system that can adapt the learning environment based on the current human LS. To design ERUDITE, we need to consider the following aspects.

1) Different mental states can affect human learning ability, such as drowsiness and stress [54], [55], [56].
2) Adapting the learning environment can have different effects on different humans. For example, some humans can experience dizziness and
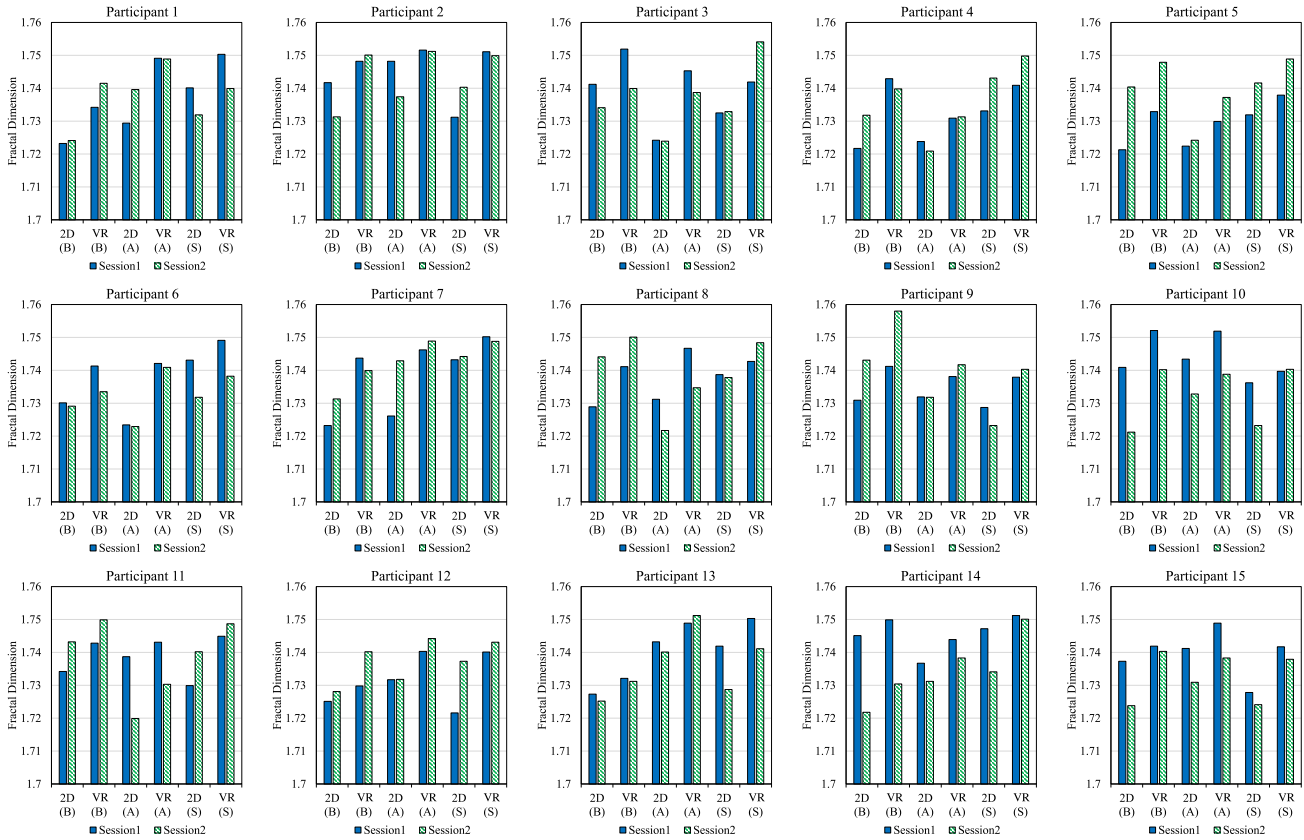
Fig. 8. Fractal dimension of EEG signal of 15 participants exposed to two presentation modalities (traditional 2-D and immersive VR). Each chart represents a participant and two sessions are depicted for each participant. B: Biology, A: Modern architecture, and S: Space.

TABLE V
SPECTRAL POWER OF THE HIGH-FREQUENCY SUB-BANDS (10–25 HZ) OF THE EEG SIGNAL FOR TWO PRESENTATION MODALITIES IN THE EXPLANATION-BASED CONCEPT LEARNING EXPERIMENT. EACH COLUMN IS AVERAGED ON THREE VIDEO CONTENT, INCLUDING BIOLOGY, MODERN ARCHITECTURE, AND SPACE

| | Traditional 2D | | Immersive VR | |
| | Session 1 | Session 2 | Session 1 | Session 2 |
Participant	$(v^2 Hz^{-1})$	$(v^2 Hz^{-1})$	$(v^2 Hz^{-1})$	$(v^2 Hz^{-1})$
Participant 1	0.075	0.059	0.102	0.102
Participant 2	0.078	0.082	0.111	0.129
Participant 3	0.069	0.071	0.130	0.119
Participant 4	0.087	0.073	0.121	0.131
Participant 5	0.081	0.097	0.118	0.129
Participant 6	0.083	0.065	0.133	0.102
Participant 7	0.065	0.062	0.102	0.095
Participant 8	0.089	0.075	0.121	0.114
Participant 9	0.084	0.091	0.121	0.126
Participant 10	0.087	0.083	0.119	0.131
Participant 11	0.091	0.088	0.112	0.136
Participant 12	0.061	0.077	0.113	0.125
Participant 13	0.083	0.089	0.141	0.127
Participant 14	0.069	0.095	0.110	0.130
Participant 15	0.091	0.097	0.118	0.129
Average	0.079	0.081	0.118	0.122

simulator sickness if exposed to an elongated VR environment [57].

3) To achieve the goal of a human-in-the-loop IoT learning system, the design of ERUDITE needs to be scalable and easy to deploy on edge devices.

A. IoT System Design

The goal of ERUDITE is to provide personalized adaptation actions in the learning environment that best fit the human to enhance the learning experience. In particular, the current human state will be used as an input to an adaptation engine to determine the correct adaptation actions. These personalized adaptations will be tuned based on the human state while interacting in the learning environment. Accordingly, there are 4 essential components for ERUDITE. First, we need to infer the human state in the learning environment. Afterward, we design an adaptation engine to provide the correct adaptation actions. Then, we need to be able to assess these adaptation actions if they enhance the learning experience. Ultimately, we need to ensure that ERUDITE computation model can be scalable and easily deployed on edge devices.

Fig. 9 illustrates ERUDITE framework. Below we will explain the different components of this framework in detail.

B. Human State in Learning Environment

Humans can lose focus or get drowsy during elongated teaching or training periods, especially in online or remote learning environments [58], which can decline human learning performance. Moreover, as we incorporate this new reality in the post-COVID-19 era, new technologies are being adopted in learning environments. In particular, while VR/mixed reality (VR/MR) technologies have been used heavily in the gaming and entertainment industry [59], recent years showed that
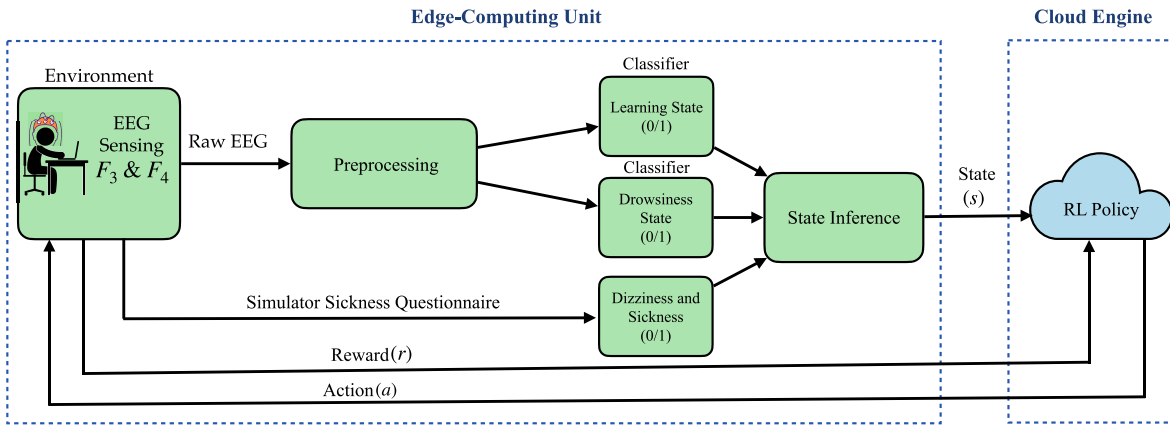
Fig. 9. ERUDITE framework for learning IoT system. EEG data is collected using a wearable device to infer the human state (s). The human state is determined through three features; the LS, the DS, and the simulator sickness scale (SSQ). The current human state s is used as an input to the ERUDITE RL engine, which selects an adaptation action a to enhance the human learning experience. This adaptation action is tuned based on a feedback reward (r) received from the human-in-the-loop environment.

these new technologies would have a significant impact on the learning [60], working and training [61] sectors. Accordingly, we exploit these technologies during learning (as we showed in Section IV). However, humans react differently to the VR environment. Some humans report dizziness and cybersickness symptoms during exposure to VR [62], which can affect the learning experience.

In ERUDITE, we focus on the current human LS, the drowsiness state (DS), and the dizziness state to infer the current human state.

1) Learning State: Based on the insights from rule-based and explanation-based concept learning, we concluded that the event of human learning correlates positively with the high-frequency sub-bands (10–25 Hz) as discussed in Sections III and IV-C2. To infer the human LS at run-time (LS), we use the DSTCLN classification model described in Section III-D to classify EEG signal into two classes of learning (1) and not-learning (0). In ERUDITE, we designed the experiment stage to be 10 min. To measure the LS for each stage, we segment the 10 min of the EEG signal into 4 s windows. Then, we estimate the LS for each of 4s windows and use the majority voting among the 4s windows to determine the LS for the duration of 10 min.

We collected the EEG data from $F3$ and $F4$ channels using the same EMOTIV_X device setup (as described in Table III) and averaged them for LS inference. Accordingly, we can then measure EEG data at run-time and use it as a sensor modality to infer the human LS.

2) Drowsiness State: While different approaches in the literature used several sensor modalities to detect drowsiness levels, such as heart rate, respiration rate, and eyelid movement [63], [64], we can exploit the same EEG signal ($F3$ and $F4$ channels) that we record to infer the DS. To quantify the DS using the EEG signal, we designed another DSTCLN model [46] to classify the EEG signal into two classes; drowsy (0) and alert (1). Details of the DSTCLN model are provided in Appendix B.

3) Cybersickness and Dizziness: The most commonly reported measure of cybersickness symptoms is the simulator

TABLE VI
CALCULATIONS IN THE SSQ. TOTAL IS THE SUM OBTAINED BY ADDING THE SYMPTOMS SCORES. OMITTED SCORES ARE ZERO

Symptoms	Subfactor		
	Nausea (N)	Oculomotor (O)	Disorientation (D)
General discomfort	1	1	
Fatigue		1	
Headache		1	
Eye strain		1	
Difficulty focusing		1	1
Increased salivation	1		
Sweating	1		
Nausea	1		1
Concentrating	1	1	
Fullness of head			1
Blurred vision		1	1
Dizzy (eyes open)			1
Dizzy (eyes closed)			1
Vertigo			1
Stomach awareness	1		
Burping	1		
Subfactor Total	N_T	O_T	D_T

sickness questionnaire (SSQ). The SSQ was derived directly from the Pensacola motion sickness questionnaire (MSQ) [65]. The MSQ consists of a list of 25 to 30 symptoms, such as spinning, tired/fatigued, and may vomit. Symptoms severity are rated on four levels, "none" (0), "slight" (1), "moderate" (2), and "severe" (3). A total score was computed by summing item scores. The highest score was determined to specify emesis as the worst case of sickness. In particular, the SSQ is a selection of 16-items as in Table VI from the MSQ with a different scoring scheme. Based on four main subfactors of cybersickness, including Nausea (N), Oculomotor (O), and Disorientation (D), related scores for the symptoms for the specific subfactor are calculated as shown in (1). Afterward, a Total Score (TS) is computed, representing the overall severity of cybersickness experienced by the users of VR systems.

Each subfactor is scaled to have a standard deviation of 15 for all observations. Subfactor scores can range from 0 to 30.54 (N), 28.58 (O), and 34.92 (D). A score of total severity (TS) is derived by summing the raw (unscaled) subfactor scores, then multiplying that sum by 3.74 as in (1). In particular, TS can range from 0 to 235.62 [66]

$$N = N\_T \times 9.54$$

$$O = O\_T \times 7.58$$
$$D = D\_T \times 13.92$$
$$\text{TS} = (N\_T + O\_T + D\_T) \times 3.74. \tag{1}$$

We use the reported value of TS to monitor the humans' simulator sickness level. In particular, participants are asked to fill out the SSQ questionnaire based on Table VI. Using this questionnaire, we calculate the TS based on (1). The threshold for TS to consider it a simulator sickness depends on the application [66].

It is recommend to choose δ_{SSQ} to be $(\text{TS}_{\max}/4)$ where $\text{TS}_{\max} = 235.62$ as mentioned above [66]. If the measured TS is bigger than the δ_{SSQ} the SSQ state is classified as dizzy (0), and if it is less than δ_{SSQ} the SSQ state is classified as is not-dizzy (1). At the end of each experiment stage, participants are asked to fill the SSQ based on Table VI. The result of the questionnaire is used to calculate the TS based on (1).

C. ERUDITE Adaptation Engine

The human state (s), a tuple of (LS, DS, SSQ), is used to choose the best adaptation action to enhance the learning experience for the human. For example, if the human is drowsy and the LS is decreasing, then one possible action could be to change the presentation modality to VR. However, this can cause dizziness/simulation sickness, as we explained before. Hence, another possible action could be to switch back to the traditional 2-D presentation or give a break. However, every human may react differently, and there are intrinsic inter- and intra-human variability. In particular, the same human within the same state can prefer an adaptation action at one time and prefer another adaptation action at another time (intrahuman variability). Moreover, different humans can have other preferences even if they are within the same state (interhuman variability) [17], [19], [20]. Accordingly, the adaptation engine has to monitor the human's preference and the effect/response of the adaptation action on the learning experience to provide personalized adaptation actions. This state-action-response interaction to tune the adaptation engine fits perfectly within the reinforcement learning (RL) paradigm. In particular, in the RL paradigm, an agent interacts with an environment by observing the state of the environment and applying an adaptation action. The RL agent then learns if this action is a good or bad action through a notion of a reward. If the agent chooses a wrong action at a particular state, the agent receives a negative reward from the environment. In contrast, if the RL agent chooses a good action, the agent receives a positive reward. Through this interaction, the RL agent converges into a policy that determines the best action per state.

By designing this RL agent, we will answer the questions (*Q2* and *Q3*) we posed in Section I.

1) Reinforcement Learning: Due to the intrinsic variability in modeling the human preferences to particular adaptation actions, the *Q*-learning algorithm can be used to model this uncertainty. Based on the current state s, Q-learning chooses an action a. Learning the optimal policy $\pi(s, a)$ —action per state that maximizes the total reward–, by applying an action

TABLE VII
HUMAN MENTAL STATE IN ERUDITE IS ONE OF EIGHT STATES DEPENDING ON THE BINARY CLASSIFICATION OF THE LS, DS, AND SIMULATOR SICKNESS (SSQ)

LS	1	1	1	1	0	0	0	0
DS	1	1	0	0	1	1	0	0
SSQ	1	0	1	0	1	0	1	0
State	s_8	s_7	s_6	s_5	s_4	s_3	s_2	s_1

(*a*) in a particular state (s) and observing the next state (s'), the RL converges to the optimal policy that maximizes the total reward. With every interaction, the RL agent updates a value for every state-action pair ($Q(s, a)$) and receives a reward ($r(s, a)$) as follows:

$$Q(s, a) \leftarrow Q(s, a) + \alpha \Big[r(s, a) + \gamma \max_a Q(s', a) - Q(s, a) \Big]. \tag{2}$$

The hyperparameters γ and α are known as the discount factor and the learning step size, respectively. To choose an action, a at each state s from the possible set of actions, an ϵ-greedy policy can be adopted. We discuss the details of hyperparameter selection in Appendix C.

2) State Space: In ERUDITE, human state refers to a combination of 3 binary features, the LS, the DS, and the simulator sickness score (SSQ). We chose binary features to reduce our state space. In particular, the LS is classified as learning (1) versus not-learning (0), the DS is classified as alert (1) versus drowsy (0), and SSQ is classified as not-dizzy (1) versus dizzy (0). Table VII shows the state space that we consider in ERUDITE. Accordingly, the best human state is s_8, where the LS of the human is high, the human is alert, and is not experiencing cybersickness. In contrast, the worst human state is s_1, where the human LS is low and is experiencing drowsiness and cybersickness. Indeed, humans can transition between any of these states.

3) Action Space: Based on the current human state, ERUDITE takes appropriate action to enhance the human learning experience. The action space in ERUDITE includes the following five actions.

1) *a_1:* Give a break to the human.
2) *a_2:* Enable VR mode by switching from 2-D to 3-D.[7]
3) *a_3:* Disable VR mode by switching from 3-D to 2-D.
4) *a_4:* Changing the content of the presentation.
5) *a_5:* No change to the learning environment.

In particular, enabling the VR mode increases brain engagement and enhances learning performance, as shown in Section IV. However, some humans may experience cyber-sickness with exposure to VR; hence ERUDITE may need to switch back to the traditional 2-D to reduce cybersickness symptoms. Moreover, a break during a learning session may also be needed to minimize drowsiness or cybersickness symptoms.

4) Reward: At the end of each experiment stage, the participant is asked to take a quiz about the presented lecture. The reward value is based on the participant's performance on the quiz and the transition state. The human performance in

---

[7]Switching from 2-D to 3-D means changing the presentation mode from 2-D video on laptop screen to a 3-D video presentation using VR device.

a quiz dictates the reward value after every learning module, where the score in this quiz is measured as a percentage. Furthermore, the next state, determined at the end of each stage of the experiment (10 min), is included in the reward as follows in (3). States with higher learning performance, lower drowsiness, and lower SSQ measures are considered better states with higher rewards. It means that state s_8 (LS = 1, DS = 1, and SSQ = 1) and state s_1 (LS = 0, DS = 0, and SSQ = 0) are the best and worst states in terms of the reward value, respectively.

Quantification of the state transition for reward is linear in the [0, 100] range with the best state allocated value 100. For instance, state s_8 and s_1 receive reward $r = 100$ and $r = 0$, respectively. For other states, reward values are distributed evenly (reward unit = 100/8). Transitions to higher states (i.e., s_6 to s_7) receive one positive unit of reward and transitions to lower states (i.e., s_6 to s_5) receive one negative unit of reward.

The quiz result is in the range of [0, 100]. Quiz quantification applies to 10 multiple-choice questions uniformly as the 10/10 and 0/10 scores receive rewards 100 and 0, respectively. Similarly to the state transitions, quiz performance improvement (i.e., score 6/10 to score 7/10) receives one unit of positive reward (+10) and a drop in quiz performance (i.e., score 8/10 to score 7/10) receives one unit of negative reward (−10)

$$r(s, a) = \text{Quiz}_{(\text{grade})} + \text{State Improvement}_{(s \rightarrow s')}. \quad (3)$$

D. ERUDITE IoT Design Choice

As we envision ERUDITE to be used in the future smart classroom, we designed ERUDITE as an edge-cloud system. In particular, for a smart classroom, ERUDITE needs to consider all students' states to ensure fairness in the classroom and to enhance system performance through scalability. Indeed, sharing the human high-level state (LS, DS, and SSQ) can be a privacy concern. However, recent work in the literature showed that even if adaptation policies run on the edge and only the adaptation action is shared with the cloud, there can still be privacy leaks through monitoring the time-series of adaptation actions [21], [67], [68]. Hence, the decision to run the RL agent in the cloud is driven by the goal of enabling the tradeoff between privacy protection, fairness of adaptation, and efficacy within ERUDITE. We discuss some of these privacy and fairness concerns in the discussion in Section VII.

VI. EVALUATION

We designed a learning environment to evaluate the proposed ERUDITE framework. We used lecture contents from Khan Academy with quizzes that cover topics on biology [69], chemistry [70], and physics [71]. We asked 15 participants (eight male and seven female) to watch these lectures, all within the age range of 20–30. The age average and standard deviation of the participants were 26 and 2.16, respectively. Each lecture is stand-alone and does not require any prior knowledge from the participants to be understood. The participants' main task was to watch the lecture and pay attention to answer the questions regarding the content at the end of the lecture. Originally, each

lecture is ≈ 55 min, in narrative style, and does not include any quizzes in the middle of them.

A. System Implementation

We used an EMOTIV EPOC_X portable EEG device (as described in Table III) to collect the EEG data from participants. Before the experiment, we present a 10-min 2-D video presentation to measure the baseline for each participant. Every 10 min, the raw EEG data are processed to measure the LS, and DS. SSQ is calculated using an online questionnaire based on Table VI. After measuring the baseline, we presented the lectures (selected from Khan Academy lectures) which were divided into 10-min videos. Each 10-min video is called one stage. Overall, 5 stages (or chunks of 10-min durations) for each lecture.[8] Between every stage, the human state is inferred as explained in Section V-B1. For each stage of the experiment (10 min), we process every 4 s and use the majority voting to infer the human state at the end of the stage. Then, the ERUDITE policy runs to determine the adequate adaptation action as explained in Section V-C3. In particular, the participant starts the video in the traditional 2-D presentation, and depending on ERUDITE policy, the learning environment can switch to VR, or the participant can take a break, or change the content, or the learning environment remains in the traditional 2-D presentation. Between every stage, ERUDITE evaluates the participant's state and applies an adaptation action.

B. Results

Table VIII shows the optimal action for each state for each participant for five stages of the experiment (each stage duration is 10 min). Each column represents the current state and each row represents the optimal action for the current state. State/action (a/s) pairs in each cell represent the action and future state (transitioned). For example, when participant 3 is in state 2 (it is highlighted in green in Table VIII), the optimal action is a_2. Taking action a_2 results in a transition from state s_2 to state s_7. In state s_2, the participant experiences low learning (LS = 0), drowsy (DS = 0), and no cybersickness (SSQ = 1) (according to the Table VII). Based on the RL model, the optimal action for state s_2 is action a_2. Action a_2 is enabling the VR mode by switching from 2-D to 3-D presentation. As the participant 3 (in the state s_2) is not experiencing cybersickness but suffering from drowsiness and a bad LS, switching to VR presentation mode is a reasonable action.

In another instance, participant 2 is in state s_6 (it is highlighted in blue in Table VIII). Optimal action in state s_6 is action a_4. Taking action a_4 results in a transition to either state s_8 or state s_7 with an equal transition probability. In state s_6 participant 2 suffers from drowsiness. VR mode is already enabled and participant 2 is comfortable with it (in the state s_6, SSQ = 1). In this situation, the RL model takes action a_4 as an optimal action which is changing the content of the lecture. By taking action a_4, in the next transitioned

---

[8]The average attention span of the human is 10 to 15 min [72].

TABLE VIII
OPTIMAL ACTION FOR DIFFERENT STATES FOR 15 PARTICIPANTS. FOR EACH STATE, EACH ROW PRESENTS THE OPTIMAL ACTION. THE EXPERIMENT
INCLUDES ONE SESSION FOR BASELINE (TO DETERMINE THE INITIAL STATE) AND THEN FIVE STAGES WHERE AT EACH STAGE,
LS, DS, AND SSQ ARE CALCULATED TO DETERMINE THE CURRENT STATE

Participant	State1	State2	State3	State4	State5	State6	State7	State8	Improvement
P1	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	21%
P2	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	27%
P3	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	24%
P4	a_1/s_8	a_1/s_8	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	24%
P5	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	26%
P6	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_1/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	19%
P7	a_1/s_8	a_2/s_7	a_3/s_8	$a_4/(s_7|s_8)$	a_3/s_8	$a_2/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	26%
P8	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	31%
P9	a_1/s_8	a_1/s_8	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	24%
P10	a_1/s_8	a_2/s_7	a_1/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	35%
P11	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	26%
P12	a_1/s_8	a_1/s_8	a_3/s_8	a_2/s_8	a_1/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	22%
P13	a_1/s_8	a_2/s_7	a_1/s_8	$a_4/(s_7|s_8)$	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	29%
P14	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_4/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	25%
P15	a_1/s_8	a_2/s_7	a_3/s_8	a_2/s_8	a_3/s_8	$a_2/(s_8|s_7)$	a_3/s_8	$a_5/(s_6|s_4)$	29%
Average	-	-	-	-	-	-	-	-	26%
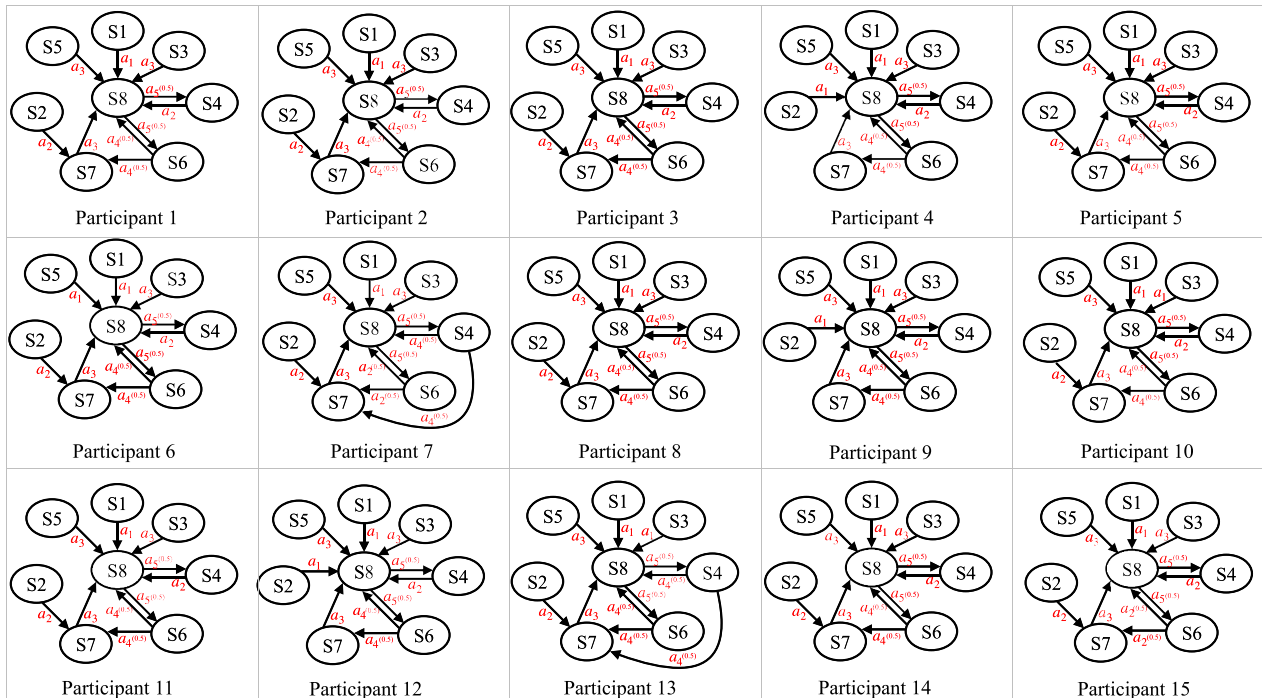
Fig. 10. Representatives of the state spaces of the participants for visualization. Each subfigure depicts the states and converged actions and their associated transition probability. Probability 1 is not displayed. Green highlight depicts the variations among the participants and shows how ERUDITE adapts itself to these variations. Each column includes some representatives for that state space.

state, either the participant 2 is yet comfortable with the VR mode (which transitions from s_6 to s_8) or it shows signs of cybersickness (which transitions from s_6 to s_7). In both transitioned states (s_7 and s_8) participant 2 is not experiencing drowsiness which was the problem in state s_6 and action a_4 aimed to address it. In the next part of the results, we show the adaptability of the ERUDITE to differences among the participants.

Fig. 10 presents state space models for representative participants in the ERUDITE evaluation for visualization. Each circle represents a state with an arrow presenting the optimal action for that state. For some actions, the transitioned state

might be more than one state. The transition probability for each action is provided in each action arrow, otherwise, the transition probability is 1. For instance, for participant 6, action a_5 transitions state s_8 to either state s_6 or state s_4 with the equal probability of 0.5.

Different humans may react differently to the same state. Fig. 10 highlights the adaption capability of the ERUDITE by presenting the differences among the state spaces of the participants in green highlights. For instance, state space for participant 6 shows that in the state s_5 the optimal action is a_1 (highlighted in green). On the other hand, the optimal action in the state s_5 for participants 4 and 9 is action a_3. This

difference depicts the adaptability of the ERUDITE to different humans.

We also quantified the improvement in the learning experience of the participants. To quantify the learning improvement, we measured the overall improvement in the quiz results and state transitions. At each stage, taken action transitions ERUDITE to the next state. The next state can be either a better state (with a positive reward) or a worse state (with a negative reward), depending on the current state. Quantification of the state transition and quiz performance is as explained in Section V-C4. For 5 stages of the experiment, 5 state transitions receive a transition reward each. Also, each stage receives a quiz performance reward. The baseline (quiz score, LS, DS, and SSQ) is compared with the new states to evaluate the ERUDITE performance. For instance, participant 1 started the experiment with the baseline state s_8 and after one stage of the experiment, it transitioned to state s_4. Transitioning from state s_8 to state s_4 receives a negative reward of -40 as explained in Section V-C4. Also, participant one's performance on the quiz decreased by 20% which receives the quiz performance reward of -20.

The last column of Table VIII illustrates the overall performance of the ERUDITE for each participant. For instance, for participant 1, the "Improvement" indicates that at the end of the experiment, the overall learning performance of this particular participant increased by 21%. This value is compared with the baseline state (state and quiz score) to calculate the improvement. As Table VIII illustrates, on average ERUDITE improves the overall performance by 26% over the participants' population.

C. Deployment and Execution Timing Analysis

The DSTCLN models were trained on a MacBook with a 2.3 GHz 8-Core Intel Core i9 M1 Processor. The trained models are then deployed and executed on an edge prototyping platform. In our evaluation, we used an ARMv8-based prototyping platform to ensure the scalability of the ERUDITE in terms of time, memory, and power consumption.

To collect the EEG signal and process it at run-time on the edge efficiently, we use Cortex API [73] which is provided by EMOTIV for brain-computer interactions and is available for macOS, Android, iOS, Windows, and various microcontrollers.

To accommodate the edge device constraints, our implementation is designed to process a 4-s window of the EEG data to infer the drowsiness and the LS at run-time from the trained DSTCLN models as explained in Sections V-B2 and III-D, respectively. The classification results of multiple 4-s windows are then collected for 10 min, then we take the majority vote as the final state inference. After we determine the LS, drowsiness, and SSQ states on the edge device (ARMv8-based prototyping platform), the tuple (LS, DS, SSQ) state is shared with the cloud that runs the RL agent to select the proper adaptation action as depicted in Fig. 9.

1) Memory Footprint: In terms of memory footprint, the trained DSTCLN models for drowsiness and LSs occupy ≈ 100 MB of memory of the ARMv8-based prototyping platform. We utilized SLOCCount [74] for analyzing significant lines of code (SLOC) for the deployed DSTCLN models (written in $C++$) on the microcontroller. It generated an output of 118 SLOC in total for both the deployed DSTCLN models for drowsiness and LSs. Moreover, as explained in Section VI-A, we only keep a 4-s window of EEG data to use for the DSTCLN models. The size of a 4-s window is approximately 2kB (256 sampling rate and 16 bits).

2) Execution Time: The execution path starts by sending the EEG data from EMOTIV to the ARMv8-based prototyping platform to infer the human state. The human state is then shared with the cloud server to take the proper action using the RL policy. As explained in Section VI-A, the EEG data for the 4-s window are passed through the DSTCLN models to infer the DS and LS. We collect these states for 10 min and then use majority voting to infer the final human state at the end of the 10-min period. This final state is the one shared with the cloud server. To ensure the responsiveness of the system, during the last 4-s window in the 10-min period, we start to share the final inferred state with the cloud server, where the ERUDITE RL policy is executed to choose the appropriate adaptation action. Hence, the last 4-s window is not used in the majority voting.

Accordingly, the main computation task that executes on the ARMv8-based prototyping platform is executing the already trained DSTCLN models, which takes approximately 1 and 1.15 s on average to infer the drowsiness and LSs, respectively. In terms of communication with the ARMv8-based prototyping platform, EMOTIV shares the EEG data via Bluetooth version 4.2 and the size of the data is approximately 2kB (256 sampling rate, 16 bits, and 4 s duration). For a Bluthooth type 4.2 it takes 16 ms to transfer 2kB of data. On the cloud server which is a MacBook with a 2.3 GHz 8-Core Intel Core i9 M1 Processor, it takes 0.12 s to run the ERUDITE RL policy to select the proper action. Since each window duration is a 4-s window, this provides enough time for the communication, EEG processing on edge-device, and RL policy on cloud server to finish their execution.

3) Power Consumption: Armv8 is engineered for low-power consumption, typically eliminating the need for heat sinks, which makes it a viable choice for edge implementations. Using ARMv8-based prototyping platform, ERUDITE consumes less than 75 mW of power on average for processing each 4-s window of EEG signal during the execution of the DSTCLN models.

VII. DISCUSSION AND FUTURE WORK

In this article, we proposed an IoT system for a human-in-the-loop learning environment. We tackled the problem of LS monitoring by exploiting two approaches of concept learning theory (rule-based learning and explanation-based learning) and showed that it is possible to infer the LS at run-time using wearable devices by decoding the EEG signal. Furthermore, we integrated this LS as a sensor modality in the learning environment to provide a personalized real-time adaptation using reinforcement learning to improve the LS of a human. Indeed, ERUDITE can be extended to address more challenges

to design an IoT learning system. Below we list our future work to extend ERUDITE.

1) False Negatives in Detecting Learning State: Concept learning perception may vary between individuals. It has been known that an individual's perception of learning influences their motivation and capability to learn. For example, if during a presentation a human believes they learned the material, it will affect their motivation and capability to learn the rest of the presentation [75]. This could signal a false negative in detecting the LS. In particular, the LS will be classified as "not-learning" while it should be "learning." In ERUDITE, we address this by using a quiz that can ensure that the human actually learned the material. Indeed, this can be extended by fusing more sensor modalities to hinder these false negatives.

2) Generalization of the Learning Measurement: To generalize ERUDITE, we need to consider the factors affecting the system. Different factors can affect human concept learning, including age and health conditions. For instance, on the one hand, younger people may comfortably use VR devices, and wearing the device may improve their concept learning performance. On the other hand, wearing VR devices by adults may increase their drowsiness and simulator sickness, leading to a worse LS. The other factor that can influence learning ability is prehealth conditions. For instance, neurodivergent populations may require different learning environments and adaptations. Considering these factors require more samples from each of these groups and studying the proper adaptation actions in the learning environment that these groups may prefer.

3) Privacy Concerns: The brain's frontal lobe is responsible for several essential tasks, such as cognitive functions, voluntary movement or activity, consciousness, memory, attention, and motivation. Collecting and decoding the EEG data from the frontal lobe (channels $F3$, Fz, and $F4$) might reveal several other critical states of the participant, which brings up the privacy problems [76], [77]. While in ERUDITE, we shared the high-level state (LS, DS, and SSQ) instead of sharing the raw EEG data with the RL adaptation engine in the cloud, sharing the high-level states, such as LS, can still be invasive. Moreover, even if the high-level states are protected, spyware monitoring the actions taken by the adaptation engine can still be used to infer the private states [67], [68]. We can extend ERUDITE, including the tradeoff between the utility of learning adaptation and the private state of the brain as part of our future work [21].

4) Fairness Concerns: In this article, we described a learning environment where the adaptation action is personalized. However, we may have multiple humans sharing the same learning environment, such as workforce training. When multiple humans share the same learning environment where one adaptation policy is applied to all of them, a fairness concern may arise. In particular, a system that monitors the aggregate learning performance across multiple humans to provide one adaptation to the learning environment may suffer from the "Matthew effect," which is summarized as the rich get richer and the poor get poorer [78]. One approach to address this is to integrate a fairness constraint in ERUDITE for the multihuman learning environment through minimizing the covariance between the learning rates across all the humans instead of considering the aggregate learning performance [17], [79].

5) WCST Limitations: While the WCST has been a valuable tool for assessing prefrontal-lobe functioning and executive functions, it is important to recognize that it may not provide as specific insights into these cognitive processes as initially believed [36], [80]. The WCST involves multiple neural circuits and may not fully capture the nuanced aspects of prefrontal-lobe functioning. Despite these limitations, we opted to use the WCST due to its established utility and the available data on its performance. Nevertheless, future research endeavors may benefit from considering alternative assessment tools and methodologies, such as Tower of Hanoi [81], [82], [83], to gain a more comprehensive understanding of prefrontal-lobe functioning and executive functions.

6) Applications Beyond Education: The potential of ERUDITE extends beyond the domain of education and personal state detection systems. Our IoT system can find applications in various systems, including neuro-recommendation systems. By adapting the concept learning theory-based approach and EEG signal decoding to different contexts, such as recommendation systems, we can enable more personalized and effective user experiences [84], [85]. Additionally, exploring the integration of ERUDITE in other personal state detection systems opens up opportunities for enhancing well-being, productivity, and user satisfaction in a wide range of applications, such as personalized medicine [86] and health monitoring [87].

VIII. CONCLUSION

Understanding how the human brain learns new concepts by decoding the EEG signals from wearable technology can open the gate to many applications in education, workforce training, and human-machine symbiosis. These applications can incorporate human learning as another sensor modality to provide a personalized learning experience. In this article, we proposed ERUDITE, a human-in-the-loop IoT learning system built upon insights from concept-learning theory and exploited the EEG signals to adapt to the learning environment. ERUDITE was evaluated across participants in a learning environment and showed that by using the brain signals as a sensor modality to infer the human LS and providing personalized adaptation to the learning environment, the participants' learning performance increased by 26.39% on average. We implemented ERUDITE on an edge-based prototype. Evaluation of hardware shows that the proposed IoT system can be implemented for devices, such as ARMv8, with a minimum RAM of 100 MB.

APPENDIX A
DSTCLN MODEL FOR LEARNING

We tuned 2 DSTCLN models for learning and drowsiness classification tasks. For learning classification, the input data is the TF image extracted from the 4s window of the EEG data. In the DSTCLN, the spectro-temporal CNN architecture

included a hierarchical CNN divided into five convolutional blocks to extract high-level features. We tuned this model and used a batch size of 16, filter size 5×5, stride size of 2, dropout ratio of 0.25, and maximum pooling. exponential linear units (ELUs) are applied as activation functions in the convolutional blocks.

APPENDIX B
DSTCLN MODEL FOR DROWSINESS

The second DSTCLN model was tuned for DS classification. The input data was composed of the TF image extracted from a $4s$ window of the EEG data. In the DSTCLN, the spectro-temporal CNN architecture included a hierarchical CNN divided into five convolutional blocks to extract high-level features. Each convolutional block includes one batch normalization layer and two convolutional layers. The final model architecture used a batch size of 32, filter size 5×5, stride size of 1, dropout ratio of 0.5, and maximum-pooling worked best. ELUs are applied as activation functions in convolutional blocks. For the rest of the architecture, we used the same Bi-LSTM network structure with 4 layers with 256 hidden units and one dropout layer. Other details of the model are explained in [46].

To train this model, we conducted an experiment and collected the EEG data of five participants watching a 2-D presentation which lasted for 50 min [88]. We used Karolinska sleepiness scale (KSS) as ground truth to label the data for the level of drowsiness [89]. KSS includes a nine-point scale ranging from 1 (extremely alert) to 9 (extremely sleepy). Before beginning the presentation, participants are asked to rate their drowsiness level based on the KSS. Every 10 min of the presentation, the participants are asked to choose from the KSS to update the drowsiness reference. The participants' labeling serves as the ground truth to infer the drowsiness level. We divided the KSS 1–6 and 7–9 as *Alert*(1) and *Drowsy*(0), respectively. Then, we used the designed DSTCLN model to classify the EEG data into these two classes.

APPENDIX C
RL HYPERPARAMETERS

In (4), hyperparameters γ and α are known as the discount factor and the learning step size, respectively

$$Q(s, a) \leftarrow Q(s, a) + \alpha \left[r(s, a) + \gamma \max_a Q(s', a) - Q(s, a) \right]. \tag{4}$$

To choose an action, a at each state s from the possible set of actions, an ϵ-greedy policy can be adopted. In the ϵ-greedy policy, the RL agent chooses the action that it believes has the best long-term effect with probability 1 - ϵ, which is the maximum value of $Q(s, a)$, and with probability ϵ, it picks an action uniformly at random. This hyperparameter ϵ (the exploration versus exploitation parameter) controls how much the RL agent is willing to explore new actions that were not taken before versus relying on the best action learned. By updating $Q(s, a)$, it is guaranteed that the optimal policy π will converge to a deterministic action a per state s that

provides the maximum reward $r(s, a)$ in a finite time steps T [90]. In this study, every time `q_value` is updated, we gradually lower ϵ following an exponential decay of 0.01.

In (4), γ determines how much the RL agent cares about rewards it receives in the distant future relative to the immediate reward. In our design, a low-discount factor $\gamma = 0.001$ is selected. Also, in (4), α (learning rate) is a hyperparameter that controls how much an agent updates its estimates of the optimal policy or value function in response to new experiences or data. A low-learning rate means the agent will be slow to adapt to new information, while a high-learning rate may result in unstable convergence or overfitting to noisy data. The optimal value of α may vary depending on the specific problem and is often determined through experimentation and tuning. In this study, we chose $\alpha = 0.05$.

ACKNOWLEDGMENT

REFERENCES

[1] O. Taiwo and A. E. Ezugwu, "Smart healthcare support for remote patient monitoring during Covid-19 quarantine," *Inform. Med. Unlocked*, vol. 20, Sep. 2020, Art. no. 100428.

[2] S. Mantena and S. Keshavjee, "Strengthening healthcare delivery with remote patient monitoring in the time of COVID-19," *BMJ Health Care Inform.*, vol. 28, no. 1, 2021, Art. no. e100302.

[3] "Size of E-learning market in 2014 and 2022." Statistica. 2022. Accessed: Feb. 15, 2023. [Online]. Available: https://www.statista.com/statistics/501104/worldwide-elearning-market-size/

[4] M. Awais et al., "LSTM-based emotion detection using physiological signals: IoT framework for healthcare and distance learning in COVID-19," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 16863–16871, Dec. 2021.

[5] S. Leroy, "Why is it so hard to do my work? The challenge of attention residue when switching between work tasks," *Org. Behav. Human Decis. Processes*, vol. 109, no. 2, pp. 168–181, 2009.

[6] G. Hong, "A review of explanation in causal inference: Methods for mediation and interaction," *J. Educ. Behav. Stat.*, vol. 42, no. 4, pp. 491–495, 2017.

[7] W. Thalheimer. "The learning-transfer evaluation model: Sending messages to enable learning effectiveness." 2018. Accessed: Feb. 15, 2023. [Online]. Available: https://WorkLearning.com/Catalog/

[8] S. Elmalaki, L. Wanner, and M. Srivastava, "CAreDroid: Adaptation framework for android context-aware applications," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 386–399.

[9] M. Taherisadr, B. Demirel, M. A. Al Faruque, and S. Elmalaki, "Future of smart classroom in the era of wearable neurotechnology," 2021, arXiv:2110.11475.

[10] X. Zhang, L. Yao, S. Zhang, S. Kanhere, M. Sheng, and Y. Liu, "Internet of Things meets brain–computer interface: A unified deep learning framework for enabling human-thing cognitive interactivity," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2084–2092, Apr. 2019.

[11] J. Zheng et al., "Multiplexing of theta and alpha rhythms in the amygdala-hippocampal circuit supports pattern separation of emotional information," *Neuron*, vol. 102, no. 4, pp. 887–898, 2019.

[12] C. Berka et al., "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, Space, Environ. Med.*, vol. 78, no. 5, pp. B231–B244, 2007.

[13] J. J. Yoo et al., "When the brain is prepared to learn: Enhancing human learning using real-time fMRI," *Neuroimage*, vol. 59, no. 1, pp. 846–852, 2012.

[14] M. N. Giannakos, K. Sharma, S. Papavlasopoulou, I. O. Pappas, and V. Kostakos, "Fitbit for learning: Towards capturing the learning experience using wearable sensing," *Int. J. Human-Comput. Stud.*, vol. 136, Apr. 2020, Art. no. 102384.

[15] N. Gao, W. Shao, M. S. Rahaman, and F. D. Salim, "n-Gage: Predicting in-class emotional, behavioural and cognitive engagement in the wild," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 1–26, 2020.

[16] J. C. Wilson, S. Nair, S. Scielzo, and E. C. Larson, "Objective measures of cognitive load using deep multi-modal learning: A use-case in aviation," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 1, pp. 1–35, 2021.

[17] S. Elmalaki, "FaiR-IoT: Fairness-aware human-in-the-loop reinforcement learning for harnessing human variability in personalized IoT," in *Proc. Int. Conf. Internet-of-Things Design Implement.*, 2021, pp. 119–132.

[18] S. Elmalaki, B. U. Demirel, M. Taherisadr, S. Stern-Nezer, J. J. Lin, and M. A. Al Faruque, "Towards Internet-of-Things for wearable neurotechnology," in *Proc. 22nd Int. Symp. Qual. Electron. Design (ISQED)*, 2021, pp. 559–565.

[19] S. Elmalaki, H.-R. Tsai, and M. Srivastava, "Sentio: Driver-in-the-loop forward collision warning using multisample reinforcement learning," in *Proc. 16th ACM Conf. Embedded Netw. Sens. Syst.*, 2018, pp. 28–40.

[20] S. Elmalaki, "MAConAuto: Framework for mobile-assisted human-in-the-loop automotive system," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2022, pp. 740–749.

[21] M. Taherisadr, S. Elmalaki, and S. Stavroulakis, "adaPARL: Adaptive privacy-aware reinforcement learning for sequential-decision making human-in-the-loop systems," in *Proc. 8th ACM/IEEE Conf. Internet Things Design Implement. (IoTDI)*, 2023, pp. 262–274.

[22] Y. Ozdem-Yilmaz and K. Bilican, "Discovery learning—Jerome Bruner," in *Science Education in Theory and Practice*. Cham, Switzerland: Springer, 2020, pp. 177–190.

[23] N. D. Goodman, J. B. Tenenbaum, J. Feldman, and T. L. Griffiths, "A rational analysis of rule-based concept learning," *Cogn. Sci.*, vol. 32, no. 1, pp. 108–154, 2008.

[24] G. DeJong and R. Mooney, "Explanation-based learning: An alternative view," *Mach. Learn.*, vol. 1, no. 2, pp. 145–176, 1986.

[25] I. Toni, M. Krams, R. Turner, and R. E. Passingham, "The time course of changes during motor sequence learning: A whole-brain fMRI study," *Neuroimage*, vol. 8, no. 1, pp. 50–61, 1998.

[26] M. H. Babini, V. V. Kulish, and H. Namazi, "Physiological state and learning ability of students in normal and virtual reality conditions: Complexity-based analysis," *J. Med. Internet Res.*, vol. 22, no. 6, 2020, Art. no. e17945.

[27] J. Xu and B. Zhong, "Review on portable EEG technology in educational research," *Comput. Human Behav.*, vol. 81, pp. 340–349, Apr. 2018.

[28] L. Chamba-Eras and J. Aguilar, "Augmented reality in a smart classroom—Case study: SaCI," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, vol. 12, no. 4, pp. 165–172, Nov. 2017.

[29] K. Ahuja et al., "EduSense: Practical classroom sensing at scale," in *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–26, 2019.

[30] M. Kwet and P. Prinsloo, "The 'smart' classroom: A new frontier in the age of the smart university," *Teach. High. Educ.*, vol. 25, no. 4, pp. 510–526, 2020.

[31] T. Sutjarittham, H. H. Gharakheili, S. S. Kanhere, and V. Sivaraman, "Experiences with IoT and AI in a smart campus for optimizing classroom usage," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7595–7607, Oct. 2019.

[32] D. A. Grant and E. A. Berg, *Wisconsin Card Sorting Test (WCST)*, PAR, Lutz, FL, USA, 1993.

[33] P. Eling, K. Derckx, and R. Maes, "On the historical and conceptual background of the Wisconsin card sorting test," *Brain Cogn.*, vol. 67, no. 3, pp. 247–253, 2008.

[34] S. Ozonoff, "Reliability and validity of the Wisconsin card sorting test in studies of autism," *Neuropsychology*, vol. 9, no. 4, p. 491, 1995.

[35] B. Kopp, F. Lange, and A. Steinke, "The reliability of the Wisconsin card sorting test in clinical practice," *Assessment*, vol. 28, no. 1, pp. 248–263, 2021.

[36] S. Miles, C. A. Howlett, C. Berryman, M. Nedeljkovic, G. L. Moseley, and A. Phillipou, "Considerations for using the Wisconsin card sorting test to assess cognitive flexibility," *Behav. Res. Methods*, vol. 53, no. 5, pp. 2083–2091, 2021.

[37] R. K. Heaton and P. Staff, *Wisconsin Card Sorting Test: Computer Version 2*, vol. 4, Psychol. Assess. Resour., Odessa, FL, USA, 1993, pp. 1–4.

[38] "Biopack." 2020. Accessed: Feb. 15, 2023. [Online]. Available: https://www.biopac.com/

[39] C. Chayer and M. Freedman, "Frontal lobe functions," *Current Neurol. Neurosci. Rep.*, vol. 1, no. 6, pp. 547–552, 2001.

[40] D. V. de Alwis and C. H. Kon, "A new way to use the Ishihara test," *J. Neurol.*, vol. 239, no. 8, pp. 451–454, 1992.

[41] Y. Chen, S. Xue, D. Li, and X. Geng, "The application of independent component analysis in removing the noise of EEG signal," in *Proc. 6th Int. Conf. Smart Grid Electr. Autom. (ICSGEA)*, 2021, pp. 138–141.

[42] S. Agarwal and M. Zubair, "Classification of alcoholic and non-alcoholic EEG signals based on sliding-SSA and independent component analysis," *IEEE Sensors J.*, vol. 21, no. 23, pp. 26198–26206, Dec. 2021.

[43] B. Boashash, *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. Amsterdam, The Netherlands: Academic, 2015.

[44] F. Pedersen, "Joint time frequency analysis in digital signal processing," 1997.

[45] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-based gait recognition using convolutional neural networks and multi-sensor fusion," *Sensors*, vol. 17, no. 12, p. 2735, 2017.

[46] J.-H. Jeong, B.-W. Yu, D.-H. Lee, and S.-W. Lee, "Classification of drowsiness levels based on a deep spatio-temporal convolutional bidirectional LSTM network using electroencephalography signals," *Brain Sci.*, vol. 9, no. 12, p. 348, 2019.

[47] "Video converter." Ani3D. 2022. Accessed: Feb. 15, 2023. [Online]. Available: https://videoconverter.wondershare.com/

[48] M. Siepmann and W. Kirch, "Effects of caffeine on topographic quantitative EEG," *Neuropsychobiology*, vol. 45, no. 3, pp. 161–166, 2002.

[49] Muscle Contraction. *Biology Muscle Contraction*. (2017). Accessed: Feb. 15, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=sZuy356qkPM

[50] ARTiculations. *When Did Modern Architecture Actually Begin?* (2018). Accessed: Feb. 15, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=Ym2CGp69oBQ

[51] I. Curiosity. *What is Our Place in the Milky Way?* 2020. Accessed: Feb. 15, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=aOPP4OV3mlI

[52] "EMOTIV EPOC+ 14 channel mobile brainwear," EMOTIV. 2020. Accessed: Feb. 15, 2023. [Online]. Available: https://www.emotiv.com/product/emotiv-epoc-14-channel-mobile-eeg/

[53] K. Foroutan-Pour, P. Dutilleul, and D. L. Smith, "Advances in the implementation of the box-counting method of fractal dimension estimation," *Appl. Math. Comput.*, vol. 105, nos. 2–3, pp. 195–210, 1999.

[54] J. A. Owens, "Sleep loss and fatigue in medical training," *Current Opin. Pulmonary Med.*, vol. 7, no. 6, pp. 411–418, 2001.

[55] K.-A. Hwang and C.-H. Yang, "Automated inattention and fatigue detection system in distance education for elementary school students," *J. Educ. Technol. Soc.*, vol. 12, no. 2, pp. 22–35, 2009.

[56] U. Ha, S. Madani, and F. Adib, "WiStress: Contactless stress monitoring using wireless signals," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 3, pp. 1–37, 2021.

[57] E. M. Kolasinski, *Simulator Sickness in Virtual Environments*, U.S. Army Res. Inst. Behav. Social Sci., Monterey, CA, USA, 1995.

[58] S. Terai et al., "Detecting learner drowsiness based on facial expressions and head movements in online courses," in *Proc. 25th Int. Conf. Intell. User Interfaces Compan.*, 2020, pp. 124–125.

[59] M. Baldauf and P. Fröhlich, "The augmented video wall: Multi-user AR interaction with public displays," in *Proc. Extended Abstracts on Human Factors Comput. Syst.*, 2013, pp. 3015–3018.

[60] M. B. Ibáñez, Á. Di Serio, D. Villarán, and C. D. Kloos, "Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness," *Comput. Educ.*, vol. 71, pp. 1–13, Feb. 2014.

[61] J. Irizarry, M. Gheisari, G. Williams, and B. N. Walker, "InfoSPOT: A mobile augmented reality method for accessing building information through a situation awareness approach," *Autom. Construct.*, vol. 33, pp. 11–23, Aug. 2013.

[62] K. Brunnström, E. Dima, T. Qureshi, M. Johanson, M. Andersson, and M. Sjöström, "Latency impact on quality of experience in a virtual reality simulator for remote control of machines," *Signal Process. Image Commun.*, vol. 89, Nov. 2020, Art. no. 116005.

[63] J. Solaz et al., "Drowsiness detection based on the analysis of breathing rate obtained from real-time image recognition," *Transp. Res. Procedia*, vol. 14, pp. 3867–3876, Jun. 2016.

[64] S. Rostaminia, A. Mayberry, D. Ganesan, B. Marlin, and J. Gummeson, "iLid: Low-power sensing of fatigue and drowsiness measures on a computational eyeglass," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 2, pp. 1–26, 2017.

[65] J. F. Golding, "Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness," *Brain Res. Bull.*, vol. 47, no. 5, pp. 507–516, 1998.

[66] W. B. Stone, III, "Psychometric evaluation of the simulator sickness questionnaire as a measure of cybersickness," Ph.D. dissertation, Dept. Psychol., Iowa State Univ., Ames, IA, USA, 2017.

[67] S. Elmalaki, B. Ho, M. Alzantot, Y. Shoukry, and M. Srivastava, "SpyCon: Adaptation based spyware in human-in-the-loop IoT," in *Proc. IEEE Security Privacy Workshops (SPW)*, 2019, pp. 163–168.

[68] S. Elmalaki, M. Alzantot, Y. Shoukry, and M. Srivastava, "VindiCo: Privacy safeguard against adaptation based spyware in human-in-the-loop IoT," 2022, *arXiv:2202.01348*.

[69] "AP college biology." KhanAcademy. 2021. Accessed: Feb. 15, 2023. [Online]. Available: https://www.khanacademy.org/science/ap-biology

[70] A. Khan. "Organic chemistry." 2021. Accessed: Feb. 15, 2023. [Online]. Available: https://www.khanacademy.org/science/organic-chemistry

[71] "AP college physics." KhanAcademy. 2021. Accessed: Feb. 15, 2023. [Online]. Available: https://www.khanacademy.org/science/ap-physics-2

[72] W. J. McKeachie and M. Svinicki, *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*, 12th ed. Boston, MA, USA: Houghton-Mifflin, 2006.

[73] "Cortex API." EMOTIV. 2022. Accessed: Feb. 15, 2023. [Online]. Available: https://emotiv.gitbook.io/cortex-api/bci

[74] "SLOCCount API." SLOCCount. 2023. Accessed: Oct. 1, 2023. [Online]. Available: https://dwheeler.com/sloccount/

[75] L. Postman, "Perception and learning," in *Psychology: A Study of a Science. Study II: Empirical Substructure and Relations with Other Sciences. The Process Areas, the Person, and Some Applied Fields: Their Place in Psychology and in Science*, S. Koch, Ed., vol. 5. New York, NY, USA: McGraw-Hill, 1962, pp. 30–113.

[76] E. Debie, N. Moustafa, and M. T. Whitty, "A privacy-preserving generative adversarial network method for securing eeg brain signals," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2020, pp. 1–8.

[77] Y. Xiao, Y. Jia, X. Cheng, J. Yu, Z. Liang, and Z. Tian, "I can see your brain: Investigating home-use electroencephalography system security," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6681–6691, Aug. 2019.

[78] T. Bol, M. de Vaan, and A. van de Rijt, "The Matthew effect in science funding," *Proc. Nat. Acad. Sci.*, vol. 115, no. 19, pp. 4887–4890, 2018.

[79] T. Zhao, M. Taherisadr, and S. Elmalaki, "FAIRO: Fairness-aware adaptation in sequential-decision making for human-in-the-loop systems," 2023, *arXiv:2307.05857*.

[80] A. Poreh, D. Pastel, A. Miller, and J. Levin, "The Cleveland sorting test: A preliminary study of an alternate form of the Wisconsin card-sorting test," *Appl. Neuropsychol. Adult*, vol. 19, no. 2, pp. 147–152, 2012.

[81] A. Kim, "Performance of the tower of hanoi and its relationship to the prefrontal cortex activity," Ph.D. dissertation, Univ. Delaware, Newark, DE, USA, 2016.

[82] M. Bustini, P. Stratta, E. Daneluzzo, R. Pollice, P. Prosperini, and A. Rossi, "Tower of hanoi and WCST performance in schizophrenia: Problem-solving capacity and clinical correlates," *J. Psychiat. Res.*, vol. 33, no. 3, pp. 285–290, 1999.

[83] K. Mitani, N. Rathnayake, U. Rathnayake, T. L. Dang, and Y. Hoshino, "Brain activity associated with the planning process during the long-time learning of the Tower of Hanoi (ToH) task: A pilot study," *Sensors*, vol. 22, no. 21, p. 8283, 2022.

[84] D. Panda, D. D. Chakladar, S. Rana, and S. Parayitam, "An EEG-based neuro-recommendation system for improving consumer purchase experience," *J. of Consum. Behav.*, to be published.

[85] H.-Y. Chang, S.-C. Huang, and J.-H. Wu, "A personalized music recommendation system based on electroencephalography feedback," *Multimedia Tools Appl.*, vol. 76, pp. 19523–19542, Oct. 2017.

[86] M. Arns, "EEG-based personalized medicine in ADHD: Individual alpha peak frequency as an endophenotype associated with nonresponse," *J. Neurother.*, vol. 16, no. 2, pp. 123–141, 2012.

[87] G. Peng, M. Nourani, J. Harvey, and H. Dave, "Personalized EEG feature selection for low-complexity seizure monitoring," *Int. J. Neural Syst.*, vol. 31, no. 08, 2021, Art. no. 2150018.

[88] ATV. *Explaining Human's Brain—Full Documentary*. (2018). Accessed: Feb. 15, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=nYiSFKlxSU4

[89] A. Shahid, K. Wilkinson, S. Marcu, and C. M. Shapiro, "Karolinska sleepiness scale (KSS)," in *STOP, THAT and One Hundred Other Sleep Scales*. New York, NY, USA: Springer, 2011, pp. 209–210.

[90] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

Mojtaba Taherisadr received the master's degree in computer engineering from the University of Central Florida, Orlando, FL, USA. He is currently pursuing the Ph.D. degree in computer engineering with the University of California at Irvine, Irvine, CA, USA.

His research is focused on the design and analysis of personalized context-aware machine learning models for healthcare systems with a focus on performance, privacy, and adaptability.

Mohammad Abdullah Al Faruque (Senior Member, IEEE) received the Ph.D. degree in computer science from Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2009.

He is currently a Full Professor with the University of California at Irvine, Irvine, CA, USA, and directing the Embedded and Cyber–Physical Systems Lab. His current research is focused on the system-level design of embedded and cyber–physical systems (CPS) with a special interest in low-power design, CPS security, and data-driven CPS design.

Dr. Al Faruque has received four best paper awards (the ACSAC 2022, the DATE 2016, the DAC 2015, and the ICCAD 2009) and many best paper award nominations. He is an IEEE CEDA Distinguished Lecturer since 2022 and an ACM Senior Member.

Salma Elmalaki received the Ph.D. degree in electrical and computer engineering from the University of California, Los Angeles, Los Angeles, CA, USA, in 2018.

She is currently an Assistant Professor with the University of California at Irvine, Irvine, CA, USA, and directing the Pervasive Autonomy Lab. Her current research is focused on cyber-physical systems with a special focus on human adaptability, fairness, and privacy guarantees.

Dr. Elmalaki is the recipient of the Best Paper Award and the Best Community Paper Award at MobiCom'15.