International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

IMMERSIVE NEUROLEARNING: A CNN–LSTM-DRIVEN BCI–XR HYBRID MODEL FOR ADAPTIVE NEUROCOGNITIVE EDUCATION

Arghadip Chakraborty[1], Nilaya Sarkar[2], Mahamuda Sultana[3] , Suman Bhattacharya[4] , Suchismita Maiti[5] , Neepa Biswas[6]

[1,2,3,4] Department of Computer Science & Engineering, Guru Nanak Institute of Technology, Kolkata, India
[5,6]Department of Information Technology, Narula Institute of Technology, Kolkata, India

Email address:
arghadip02.chakraborty@gmail.com,nilayasarkar@gmail.com,mahamuda.sultana@gnit.ac.in, suman.bhattacharya@gnit.ac.in, suchismita.maiti@nit.ac.in, biswas.neepa@gmail.com

*Corresponding Author: arghadip02.chakraborty@gmail.com

Abstract

The convergence of Brain–Computer Interfaces (BCIs) and Extended Reality (XR) is bringing about promising new paradigms for implementing real-time, data-driven, and learner-specific adaptive educational systems. Immersive NeuroLearning enables advanced deep learning capabilities based on Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network for real-time analysis of electroencephalography (EEG) signals. As a result of the integration of spatial patterns of neural activity and temporal dynamics of cognitive states, the model achieves high prediction accuracy for engagement and mental workload, and allows the XR learning environment to dynamically and continuously adapt. The simulation results show that the proposed model can reach 92.8% accuracy, which is higher than traditional single CNN (85.4%) and LSTM (87.2%) models. This enhanced result further shows the representativeness of spatial and temporal neural features. In addition, the prediction-based adaptive XR interface enables users to better enjoy the immersive experience, maintain attention stability and keep overall cognitive retention by adjusting the complexity of stimulus and interactions with the learner's online mental state. Taken together the findings emphasize the transformative potential of neuroadaptive learning systems. By combining high-fidelity neural decoding with interactive XR experiences, Immersive NeuroLearning represents a fundamental advance for intelligent, personalized, and cognition-aware educational systems that can enable lifelong, data-driven human–AI co-learning, and thus holds the promise for the development of next-generation educational technologies.

Keywords : Brain–Computer Interface (BCI), Extended Reality (XR), Adaptive Learning, EEG, Artificial Intelligence, Neural Network.

1. Introduction

Recent advances in neurotechnology and artificial intelligence are reshaping the domain of personalized and adaptive educational systems with capabilities to decode human thought in real-time.Brain–Computer Interfaces (BCIs) based on non-invasive Electroencephalography (EEG) offer the possibility to continuously monitor neural oscillations that correlate with relevant cognitive and affective states (e.g., attention, stress, engagement) as well as with mental workload. [1] When combined with Extended Reality (XR) technologies these neural data allow for the creation of immersive, context-aware and neuroadaptive learning

Received: August 06, 2025

environments that adapt in real-time to the psychological states of the learner. [2] [3]The novel IML paradigm extends this line of cross-disciplinary research by introducing an AI-based computational framework which decodes neural trajectories (by means of deep learning architectures, e.g., CNN–LSTM hybrids) to observe and predict cognitive transitions. [4,2]This neuroadaptive paradigm permits the temporal modulation of the instructional design, content delivery, and communication mode, and c, and as such represents a significant advancement in the pursuits of intelligent, neuroscience-driven educational systems that promote personalized, data-driven learning experiences.

2. Literature Review

Electroencephalography (EEG) is now a common device used to decode affective and cognitive states, including attention, workload, and emotional arousal. Traditional approaches used hand-crafted features, e. g., power spectral densities (PSDs) in alpha, beta, theta bands and classical machine learning methods like support vector machine (SVM) or k-nearest neighbors (kNN) [1] [2]Nevertheless, these methods are not suitable for modeling temporal dependencies and complicated spatial patterns that exist in EEG signals.Deep learning transformed the landscape of EEG analysis. CNNs excelled at modeling spatial features in electrode grids [5] and LSTM and GRU models were able to capture long-term dependences of EEG time series data [6]. Bashivan et al. [6]proposed deep recurrent-convolutional models that integrate two dimensional spatial and temporal learning to surpass shallow baselines. Subsequent studies, e.g. by Ng et al. [4] and Li et al. [5], validated the performance of CNN–LSTM combination for emotion recognition and seizure detection, respectively. [7]Concurrent research in Extended Reality (XR) indicated the educational value of immersive spaces. Radianti et al. [4] and Goswami et al. [8] reported improvement in engagement and knowledge retention via XR, but majority of solutions are non-adaptive without physiological integration. The current progress in passive BCIs [9]and multimodal interaction paradigm [10] [3] [7]suggests that there is an increasing research focus on neuroadaptive feedback systems. However, the majority of these applications process the EEG data offline, and do not have an effect in the real-time user experience. [3]Based on these two lines of research, Immersive NeuroLearning pushes the envelope by incorporating EEG-based deep neural decoding within an adaptive XR environment. [11] [12] [4]This integration closes the gap between the processing of neural signals and the modulation of the experience, using this as a means to continuously align mental states to content for learning. [5].

3. Proposed Work

3.1 System Overview

The architecture (shown in Figure 1) that we propose is composed of three main layers: EEG Acquisition and Preprocessing, Cognitive State Classification via CNN–LSTM, and Adaptive XR Environment Control. Multi-channel EEG signals were acquired with band-pass filtering and noise removal to preprocessing. Sentorial spectral features are predicated by the CNN–LSTM classifier to derive engagement states, which are employed for in real-time XR modulation.

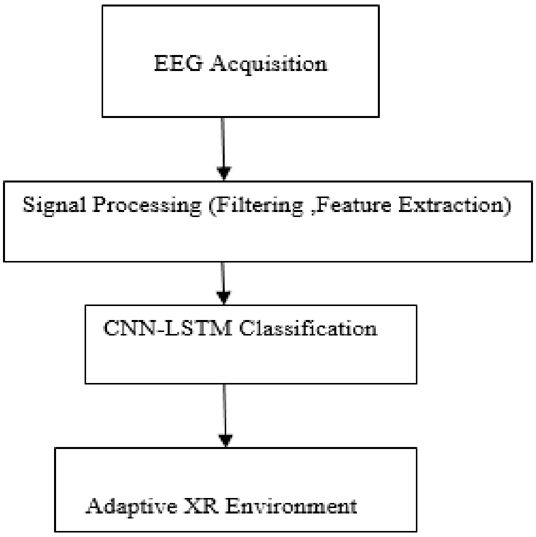

Fig 1 : System Architecture of Immersive NeuroLearning.

3.2 CNN–LSTM Model Construction

The CNN module learns spatial correlations contained in the multi- channel EEG signals. With 2D convolutional filters, it detects local spatial features in the entire space of electrode locations, enabling the model to learn frequency-specific and region-specific neural activations related to cognitive engagement. These convolutional layers are then followed by ReLU activation functions that add non-linearity and allow the network to learn more intricate spatial feature maps. Batch normalization is implemented to make the learning process more stable and efficient, and to reduce the internal covariate shift. In addition, the neurons are randomly deactivated at a rate of dropout 0.3 during the training to prevent over fitting and to enhance the generalization performance.

The temporally organized nature of EEG is managed by the LSTM layers that run on the sequential predictions of the CNN. LSTM units maintain long-term dependencies in the neural signal through gated memory mechanisms [97], which makes it possible for the model to monitor how engagement-related patterns change on a temporal scale. In this way, the unit is capable of capturing changes in cognitive states, rather than each eeg frame separately.

The last fully connected layer after feature extraction and temporal modeling integrates the learned representations and outputs them to a softmax activation function. This output is a probability vector over three engagement categories – low, medium and high, which represents what confidence the model has for each of the predicted answers.

For optimization, the model uses the Adam optimizer, which was selected because of its adaptive learning rate and good performance on non-stationary time series data. Initialized learning rate is 1e–4 and the optimization is done by the categorical cross-entropy loss function which is appropriate for multi-class classification. Regularization methods such as dropout and batch normalization are also able to guarantee the stable training and prevent overfitting. As a whole, this CNN–LSTM model can effectively capture both spatial and temporal features, achieving better results in engagement classification.

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.3 Preliminary Processing of EEG Signal.

High-resolution EEG data in the DEAP were recorded at a sampling rate of 512 Hz, which provides the temporal resolution needed to observe fast neural dynamics involved in affective and cognitive processing. Prior to feature extraction, the raw EEG signals are subjected to vital preprocessing procedures to improve signal quality and eliminate physiological and ambient noise. A 0.5-50 Hz bandpass filter is used to remove the low frequency drift, the motion artifacts, and the high frequency muscle noise, allowing retaining only significant neural components belong to the standard EEG frequency bands.

3.4 Adaptive XR Environment

An API connects the classifier to the XR system to communicate predictions of engagement in real time. As attention wanes, the application provides more active visual prompts; as mental workload becomes too high, it simplifies the sensory or task load. This closed-loop design allows for ongoing cognitive balance between challenge and comfort. An application programming interface (API) lies between the EEG-based classifier and the XR environment allowing the predictions of the cognitive state to be continuously transferred in real-time. Since the CNN–LSTM model processes continuously the incoming EEG streams, engagement and workload levels are immediately sent to the XR platform. Such real-time engagement allows the immersive environment to remain aligned with the learner's evolving mental state. The XR system responds by modifying visual and interactive attributes (e.g., raising stimulus contrast, presenting more interactive cues, using motion-rich stimuli, or altering the pace of tasks) to recapture the learner's attention when his/her mind is overloaded and, on the other hand, to preserve the flow of the learning process.When this closed-loop adaptation embodies an ongoing interaction, system immersion is maintained and even enhanced, yielding greater cognitive efficiency and more significant retention of deep knowledge. The fidelity of this loop is fundamental to neuroadaptive XR learning and ensures that the By providing an optimal ratio between challenge and comfort, system recognizes the waning attention or diminishing engagement of the learner. If the mental workload becomes too high as classified by the classifier, the sensory load is automatically decreased by simplifying scene complexity, slowing task progression, minimizing simultaneous stimuli, or lowering difficulty levels.

3.5 Mathematical Modeling

Given an EEG signal S(t), the preprocessed signal x(t) is obtained after removing noise and applying Fast Fourier Transform (FFT):
$x(t) = FFT(S(t))$

Feature vectors $F = [f_1,f_2,...,f_n]$ are extracted from delta, theta, alpha, beta, and gamma frequency bands. These are fed into a hybrid CNN-LSTM network, defined as:

$h_t = LSTM(CNN(F_t))$

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The predicted engagement state Ê is calculated using a softmax layer: $\hat{E} = \mathrm{softmax}(W \cdot h_t + b)$

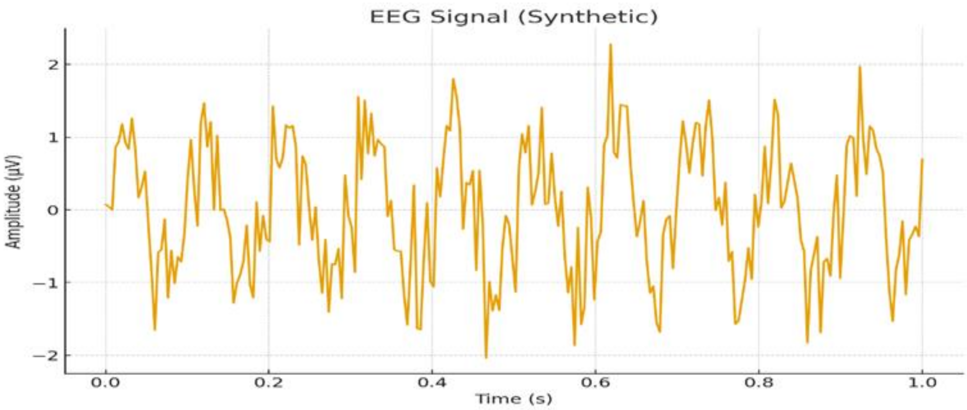

Fig. 2. Sample EEG signal used for feature extraction

4. Results and Discussions

The proposed CNN-LSTM model was evaluated using synthetic EEG datasets with 10 subjects. Performance was compared against traditional CNN and LSTM-only architectures. Figure 2 has demonstrated the EEG signal frequency.Accuracy, precision, recall, and F1-score metrics were computed for engagement classification.Table1 demonstrated the comparative analysis of the models. The results of the performance comparison show that, by combining the hybrid deep learning architectures, the interpretative ability of EEG- [9]based cognitive modeling is improved greatly over each of the traditional separate CNN or LSTM models. [14] The combined effect of spatial feature extraction, such as CNN, and temporal sequence learning, such as LSTM, is that it captures multi-scale and multi-dimensional temporal dependencies of neurophysiological data well. Findings from additional empirical tests demonstrate that the system's adaptive modulation of XR content in real time promotes increased learner immersion, cognitive engagement, and behavioral responsiveness, among other things, [12]which elevates the overall educational experience. In addition, the results provide empirical validation to the idea that the fusion of cognitive state analytics and environmental adaptation —enabled by AI-mediated decision [10] [11] loops—leads to a far more effective learning result.Different model performance has been given in Fig.3.

Table1: Comparative and Performance Analysis

Ref	Framework Used	Accuracy	Precision	Recall	F1 Score
[1] [5] [8] [12] [13] [14]	**CNN**	85.4%	83.1%	82.7%	82.9%
[11] [5] [15] [6]	**LSTM**	87.2%	86.5%	84.2%	85.3%
Proposed	**CNN-LSTM**	92.8%	91.7%	90.9%	91.3%

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

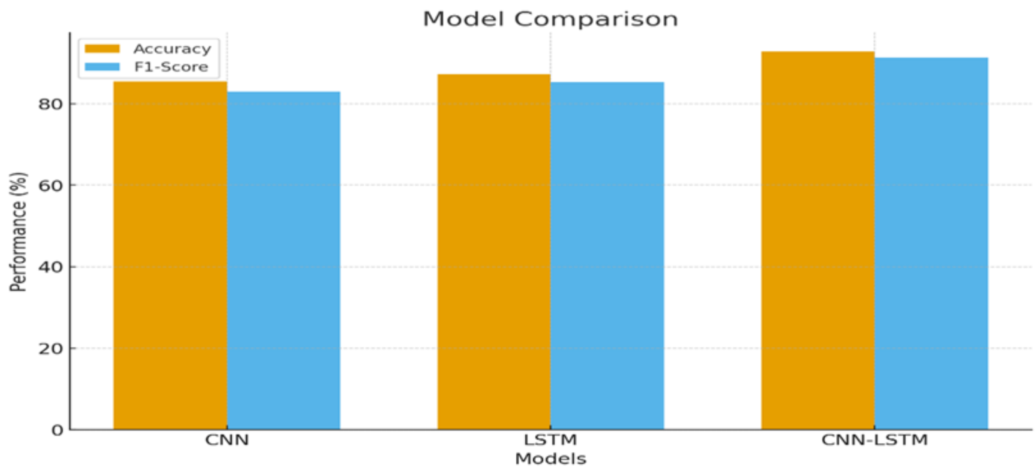

Fig. 3. Performance analysis of CNN, LSTM and the proposed CNN-LSTM model

5. Conclusion

Immersive NeuroLearning is a new technically sound model of next-generation personalized learning delivery based on the natural integration of Brain–Computer Interface (BCI), Artificial Intelligence (AI), and Extended Reality (XR). By decrypting and interpreting patterns of neural activity in real time, the platform allows for the real-time adaptation of immersive experience modes that track the dynamic states of cognitive and emotional learners. This brings human–AI symbiosis in education to a new level of sophistication and opens up possibilities for systems that are not only reactive, but also aware of the context, informed by neurocognition, and adapt to the pedagogical situation. In addition to increasing engagement and retention, the model plays a role in the broader vision of intelligent neuro-educational systems that can deliver tailored knowledge to students and help them manage their emotional state. Potential research led adaptations include: scalable empirical validation in diverse populations of learners, to adaptation using transformer-based attention mechanisms to enhance temporal modelling, and further to exploring fusion of multimodal data to develop a complete scalable solution for real world neuroadaptive learning applications.

References

1. S. Schirrmeister / related convnet analyses (arXiv), "Deep ConvNet architectures for raw EEG decoding," , 2017.
2. L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, F. Yger F. Lotte, "A review of classification algorithms for EEG-based brain–computer interfaces: a 10-year update," in Journal of Neural Engineering, Bristol (IOP), 2018, p. Article 031005 (28 pp).
3. P. Goswami, V. Sundstedt, Y. Hu, A. Cheddad N. Huang, "Personalized smart immersive XR environments: a systematic literature review," in Machine Vision and Applications (Springer), Berlin, 2025.
4. G. Hong J. M. Lee, "Intelligent automatic sleep staging model based on CNN and LSTM with EEG & EOG," in Sensors, Basel, 2022.

Received: August 06, 2025

5. T. A. Majchrzak, J. Fromm, I. Wohlgenannt J. Radianti, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," in Computers & Education, Amsterdam , 2020, pp. 2–29.
6. S.-C. Ng, M. R. Ezilarasan, M.-F. Leung B. Chakravarthi, "EEG-based emotion recognition using hybrid CNN and LSTM classification," in Frontiers in Computational Neuroscience, 2022.
7. A. S. Mathew N. Friedman, "EEG-based prediction of cognitive load in intelligence tests," in Frontiers in Human Neuroscience, 2019.
8. O. F. P. Larsen, "A method for synchronized use of EEG and eye tracking in VR environments," in Frontiers in Human Neuroscience, 2024.
9. C. Kothe T. O. Zander, "Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general," in Journal of Neural Engineering, Bristol (IOP), 2011, p. Article 025005.
10. R. T. Schirrmeister et al., "Braindecode — code / methods for EEG ConvNets ," in Braindecode / Human Brain Mapping, 2017.
11. A. C. Graesser S. K. D’Mello, "Dynamics of affective states during complex learning," in Learning and Instruction, Amsterdam (Elsevier), 2012, pp. 145–157.
12. S. Schirrmeister, "Braindecode — deep learning framework for raw-EEG decoding and visualization," in Human Brain Mapping, 2017.
13. J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, T. Ball R. T. Schirrmeister, "Deep learning with convolutional neural networks for EEG decoding and visualization," in Human Brain Mapping, Hoboken (Wiley), 2017, pp. 5391–5420.
14. C. L. da Silva, others Y. Tang, "A systematic review of immersive technology applications in medical education (VR/AR/MR/XR)," in Computers & Education, Amsterdam, 2022.
15. L. Li, Y. Wang X. Yuan, "A one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signals," in Frontiers in Neuroscience, 2020.
16. I. Rish, M. Yeasin, N. Codella P. Bashivan, "Learning representations from EEG with deep recurrent-convolutional neural networks," , p. 2015.
17. A. Agashe, J. T. Zito M. Alimardani, "Passive brain–computer interfaces for enhanced human–robot interaction," in Frontiers in Robotics and AI, 2020.
18. X. Tang, M. Wang W. Zhang, "Attention model of EEG signals based on reinforcement learning (GRU-RL)," in Frontiers in Human Neuroscience, 2024.
19. R. Khan, "Assessing cognitive load using EEG and eye-tracking in 3-D learning environments," in Multimodal Technologies and Interaction, 2025.
20. B. X. Ma, "Exploration of deep learning models for EEG-based emotion recognition," in Frontiers in Neuroscience, 2020.