JOURNAL OF ARTIFICIAL INTELLIGENCE AND ROBOTICS
Jan-Mar 2025, VOL. 2, ISSUE 1, pp. 12-14
https://doi.org/10.61577/jaiar.2025.100004

RESEAPRO JOURNALS

MINI REVIEW

OPEN ACCESS

Deep learning approaches for brain-computer interface based analysis

Jullien Holmes and Cecilia Dewandel

Department of Biomedical Engineering, Silesian University of Technology, Gliwice, Poland

ABSTRACT

The human brain exhibits distinct neurophysiological responses during deceptive and truthful states, making it a compelling target for lie detection. Traditional polygraph-based approaches suffer from reliability issues due to their dependence on physiological arousal rather than direct brain activity. Brain-Computer Interfaces (BCIs) have emerged as a promising alternative, leveraging neuroimaging techniques such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) to capture brain activity associated with deception. The integration of deep learning with BCIs has transformed the field of lie detection, offering advanced computational techniques for extracting deception-related neural signatures. Unlike conventional machine learning models, deep neural networks (DNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs) enable automated feature learning, capturing complex spatial and temporal patterns in brain signals with unprecedented accuracy. The introduction of hybrid models, including CNN-LSTM and transformer-based architectures, has further expanded the capability of these systems, allowing for more refined deception classification. Additionally, generative models such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) have been employed to address data scarcity by synthesizing realistic neuroimaging data, improving model robustness and generalization. Recent advancements in self-supervised and reinforcement learning approaches have demonstrated potential in enhancing adaptability and performance in real-world deception detection scenarios. The increasing use of multimodal fusion, integrating EEG, fMRI, and physiological data, promises to improve detection accuracy and reliability. However, several challenges remain, including ethical concerns regarding privacy, the need for large-scale, high-quality datasets, and the interpretability of deep learning models. Future research should focus on developing transparent and explainable AI systems while addressing ethical considerations to ensure responsible deployment in forensic, security, and clinical applications. The convergence of neuroscience, artificial intelligence, and cognitive science marks a transformative step toward the development of objective, data-driven deception detection technologies.

KEYWORDS

Brain-Computer Interface (BCI); Deep learning; Lie detection; Electroencephalography; Automated response

ARTICLE HISTORY

Received 06 January 2025;
Revised 27 January 2025;
Accepted 04 February 2025

Introduction

Lie detection has long been a topic of scientific inquiry, with applications spanning forensic investigations, security screenings, and clinical assessments [1]. Traditional methods, including polygraph tests, rely on physiological markers such as heart rate, blood pressure, and skin conductance [2]. However, these markers can be influenced by anxiety, stress, or countermeasures, rendering the polygraph unreliable. To overcome these limitations, researchers have explored brain activity as a more direct indicator of deception.

Brain-Computer Interfaces (BCIs) provide a novel approach to lie detection by decoding neurophysiological signals associated with cognitive and emotional states. EEG and fMRI have emerged as the primary modalities for capturing neural responses linked to deception [3]. However, conventional machine learning techniques often struggle to handle the high-dimensional, nonlinear nature of brain signals. The rise of deep learning has revolutionized this field, offering powerful tools to extract meaningful features and improve classification accuracy.

Recent advancements in computational neuroscience, particularly through deep learning models, have enabled the detection of deception-related neural patterns with unprecedented precision. Deep learning methods facilitate automated feature extraction, allowing researchers to bypass the limitations of traditional handcrafted feature engineering [4]. This review examines recent advancements in deep learning applications for BCI-based lie detection. It discusses the underlying neural mechanisms of deception, the role of EEG and fMRI in recording deception-related brain activity, and how deep learning models enhance the interpretation of these signals. Additionally, the review explores hybrid models, generative approaches, and multimodal fusion techniques that enhance the robustness and generalizability of deception detection systems. Challenges and future directions are also considered, highlighting the interdisciplinary nature of this emerging research area.

Neural Mechanisms of Deception

Deception is a complex cognitive process that involves multiple

***Correspondence:** Jullien Holmes, Department of Biomedical Engineering, Silesian University of Technology, Gliwice, Poland, e-mail: JNholmes@yaghoo.com

brain regions, including the prefrontal cortex (PFC), anterior cingulate cortex (ACC), and parietal regions. Studies using fMRI have demonstrated increased activity in the PFC during deceptive responses, reflecting the cognitive load required to suppress the truth and construct a falsehood. The ACC is implicated in conflict monitoring, detecting inconsistencies between internal knowledge and external responses. The parietal cortex, particularly the inferior parietal lobule, has been associated with attentional control and memory retrieval, which are essential during deception [5].

EEG studies have identified deception-related event-related potentials (ERPs), such as the P300 and N400 components, which signify recognition and semantic processing discrepancies. The P300 wave is particularly relevant in concealed information tests, where a suspect's recognition of critical details elicits a neural response even if they attempt to suppress it. Additionally, frequency-domain features such as increased beta and gamma band activity during deception further confirm the neurophysiological underpinnings of lying. Deep learning models can exploit these deception-related neural signatures to differentiate between truthful and deceptive responses with high accuracy. By learning hierarchical representations of neural activity, deep networks enable more precise classification of deceptive states compared to conventional machine learning methods [6]. Advanced deep learning architectures leverage these brain signal patterns to enhance accuracy and robustness in real-world scenarios.

EEG-Based Deep Learning Models for Lie Detection

EEG is the most widely used modality for BCI-based lie detection due to its high temporal resolution, non-invasiveness, and cost-effectiveness. EEG signals, however, are inherently noisy and exhibit significant inter-subject variability, making their interpretation challenging [7]. Deep learning models have been employed to address these challenges by automatically extracting relevant features from raw EEG signals.

Convolutional Neural Networks (CNNs)

CNNs have proven effective in EEG-based lie detection by capturing spatial and spectral features from multi-channel EEG recordings. These networks apply convolutional filters to extract frequency-domain information relevant to deception-related brain activity. Several studies have shown that CNN-based models can outperform traditional classifiers such as support vector machines (SVMs) and k-nearest neighbors (KNNs) in lie detection tasks [8]. Recent advancements involve the use of multi-scale CNNs, which can learn hierarchical feature representations, improving classification accuracy across diverse datasets.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) Networks

Given the temporal nature of EEG signals, RNNs and LSTMs have been employed to model time-dependent patterns associated with deception. These architectures can capture long-range dependencies in EEG sequences, improving classification accuracy. Hybrid models combining CNNs and LSTMs have further enhanced performance by integrating spatial and temporal feature extraction [9]. Additionally, attention-based RNNs have been explored to dynamically focus on the most relevant neural activity patterns, further refining lie detection accuracy.

fMRI-Based Deep Learning Models for Lie Detection

fMRI provides high spatial resolution imaging of brain activity, making it a valuable tool for identifying deception-related neural activation patterns [10]. Deep learning techniques have been applied to analyze fMRI data, extracting relevant features for deception classification.

Deep Belief Networks (DBNs) and Autoencoders

DBNs and autoencoders have been used to learn hierarchical representations from fMRI data, reducing dimensionality while preserving critical information [11]. These models enhance the detection of deception-related activation patterns, improving classification robustness. Additionally, generative adversarial networks (GANs) have been applied to synthesize realistic fMRI data to augment training datasets and address the issue of limited real-world samples.

Transformer Models for BCI-based Lie Detection

Recent advancements in transformer-based architectures, such as Vision Transformers (ViTs) and BERT-like models adapted for time-series data, have shown promise in fMRI and EEG-based deception detection [12]. Transformers leverage self-attention mechanisms to model long-range dependencies, making them well-suited for analyzing complex neural data. The application of pre-trained transformer models on large-scale neuroimaging datasets has demonstrated improved generalization and robustness, potentially paving the way for real-world deployment.

Challenges and Future Directions

Despite the success of deep learning approaches in BCI-based deception detection, several challenges persist. Data scarcity remains a significant issue, as high-quality deception datasets are limited. Transfer learning and data augmentation techniques offer potential solutions to enhance model generalization. Additionally, ethical concerns regarding privacy, consent, and potential misuse of AI-driven lie detection must be addressed through transparent guidelines and regulatory frameworks.

The interpretability of deep learning models is another challenge, as these models are often considered "black boxes." Explainable AI (XAI) techniques are needed to enhance the transparency and trustworthiness of deception detection systems. Methods such as Grad-CAM for visualizing model attention and SHAP values for feature importance analysis are being explored to improve interpretability. Future research should focus on refining deep learning methodologies while ensuring ethical and responsible deployment in forensic, security, and clinical settings.

Data scarcity and generalization

High-quality deception datasets are limited, posing challenges for training robust deep learning models [13]. Transfer learning and data augmentation techniques are potential solutions to enhance model generalization. The creation of standardized, large-scale datasets will be crucial for advancing this research field.

https://doi.org/10.61577/jaiar.2025.100004

Ethical and legal considerations

The application of AI-driven lie detection raises ethical concerns regarding privacy, consent, and potential misuse. Transparent guidelines and regulatory frameworks are necessary to ensure ethical deployment. In addition, public acceptance of AI-based lie detection remains a significant challenge, requiring extensive validation and policy discussions [14].

Interpretability and explainability

Deep learning models are often considered black boxes, making it difficult to interpret their decision-making processes. Explainable AI (XAI) techniques are needed to enhance the transparency and trustworthiness of deception detection systems. Methods such as Grad-CAM for visualizing model attention and SHAP values for feature importance analysis are being explored to improve interpretability [15].

Conclusions

The integration of deep learning with BCI technology has significantly advanced the field of lie detection. EEG and fMRI-based deep learning models demonstrate superior accuracy in identifying deception-related neural patterns compared to traditional methods. CNNs, RNNs, and transformer-based architectures have emerged as powerful tools for decoding complex brain signals, offering a data-driven approach to lie detection. However, challenges related to data availability, ethical implications, and model interpretability must be addressed to enable practical applications. Future research should focus on developing more robust, interpretable, and ethically responsible AI-driven deception detection systems, ensuring their viability for forensic, security, and clinical settings.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

1. Masip J. Deception detection: State of the art and future prospects. Psicothema. 2017;29(2):149-59. https://doi.org/10.7334/psicothema2017.34
2. Samuel SG, Chatterjee T, Thapliyal H, Kacker P. Facial psychophysiology in forensic investigation: A novel idea for deception detection. Forensic Dent Sci. 2019;11(2):90-94. https://doi.org/10.4103/jfo.jfds_49_19
3. Langleben DD, Schroeder L, Maldjian JA, Gur RC, McDonald S, Ragland JD, O'Brien CP, Childress AR. Brain activity during simulated deception: an event-related functional magnetic resonance study. Neuroimage. 2002;15(3):727-732. https://doi.org/10.1006/nimg.2001.1003
4. Çayir A, Yenidoğan I, Dağ H. Feature extraction based on deep learning for some traditional machine learning methods. In2018 3rd International conference on computer science and engineering (UBMK) 2018:494-497. IEEE. https://doi.org/10.1109/UBMK.2018.8566383
5. Christ SE, Van Essen DC, Watson JM, Brubaker LE, McDermott KB. The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. Cereb cortex. 2009;19(7):1557-1566. https://doi.org/10.1093/cercor/bhn189
6. Tang H, Lu X, Cui Z, Feng C, Lin Q, Cui X, Su S, Liu C. Resting-state functional connectivity and deception: exploring individualized deceptive propensity by machine learning. Neuroscience. 2018;395: 101-112. https://doi.org/10.1016/j.neuroscience.2018.10.036
7. Kusumawati D, Ilham AA, Achmad A, Nurtanio I. Performance Analysis of Feature Mel Frequency Cepstral Coefficient and Short Time Fourier Transform Input for Lie Detection using Convolutional Neural Network. JOIV: International Journal on Informatics Visualization. 2024;8(1):279-288. https://dx.doi.org/10.62527/joiv.8.1.2062
8. Massi MC, Ieva F. Learning Signal Representations for EEG Cross-Subject Channel Selection and Trial Classification. In2021 IEEE 31st International Workshop on Machine Learning for Signal Processing (MLSP) 2021:1-6. IEEE. https://doi.org/10.1109/MLSP52302.2021.9596522
9. Kothuri SR, Sujay V, Juyal P, Bhuvaneshwari R. Hybrid CNN-LSTM machine learning algorithm for driver distraction detection. In2024 International Conference on Integrated Circuits and Communication Systems (ICICACS) 2024:1-5. IEEE. https://doi.org/10.1109/ICICACS60521.2024.10498991
10. Cui Q, Vanman EJ, Wei D, Yang W, Jia L, Zhang Q. Detection of deception based on fMRI activation patterns underlying the production of a deceptive response and receiving feedback about the success of the deception after a mock murder crime. Soc Cogn Affect Neurosci. 2014;9(10):1472-1480. https://doi.org/10.1093/scan/nst134
11. Zhang S, Dong Q, Zhang W, Huang H, Zhu D, Liu T. Discovering hierarchical common brain networks via multimodal deep belief network. Med Image Anal. 2019;54:238-252. https://doi.org/10.1016/j.media.2019.03.011
12. Khadka R, Lind PG, Mello G, Riegler MA, Yazidi A. Inducing inductive bias in vision transformer for eeg classification. InICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2024:2096-2100. IEEE. https://doi.org/10.1109/ICASSP48485.2024.10446429
13. Wei W, Liu L. Robust deep learning ensemble against deception. IEEE Transactions on Dependable and Secure Computing. 2020; 18(4):1513-1527.0020. https://doi.org/10.1109/TDSC.2020.3024660
14. Kalodanis K, Rizomiliotis P, Feretzakis G, Papapavlou C, Anagnostopoulos D. High-Risk AI Systems—Lie Detection Application. Future Internet. 2025;17(1):26. https://doi.org/10.3390/fi17010026
15. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput Vis. 2020;128:336-359. Available at: https://openaccess.thecvf.com/content_ICCV_2017/papers/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.pdf

https://doi.org/10.61577/jaiar.2025.100004