Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 324-339 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Optimized Metaheuristic Integrated Neuro-Fuzzy Deep Learning Framework for EEG-Based Lie Detection

Tanmayi Nagale[1], and Anand Khandare[2]

[1] Department of Computer Engineering, Thakur College of Engineering and Technology, Kandivali, Mumbai, India
[2] Department of AI & DS, Thakur College of Engineering and Technology, Kandivali, Mumbai, India

Corresponding author: Tanmayi Nagale. (e-mail: tanmayinagale13@gmail.com), Dr. Anand Khandare (e-mail: anand.khandare@thakureducation.org)

Abstract **EEG-based deception detection remains challenging due to three critical limitations: high inter-subject variability, which restricts generalization, the black-box nature of deep learning models that undermines forensic interpretability, and substantial computational overhead arising from high-dimensional multi-channel EEG data. Although recent state-of-the-art approaches report accuracies of 82–88%, they fail to provide the transparency required for legal and forensic admissibility. To address these limitations, this study aims to develop an accurate, computationally efficient, and explainable EEG-based deception detection framework suitable for real-world forensic applications. The primary contribution of this work is a novel hybrid neuro-fuzzy architecture that jointly integrates intelligent channel selection, complementary deep feature learning, and transparent fuzzy reasoning, enabling high performance without sacrificing interpretability. The proposed framework follows a five-stage pipeline: (1) intelligent channel selection using Type-2 fuzzy inference with ANFIS-based ranking and multi-objective evolutionary optimization (MOEA/D), reducing EEG dimensionality from 64 to 14 channels (78.1% reduction); (2) dual-path deep learning that combines EEGNet for spatial–temporal feature extraction with InceptionTime-Light for multi-scale temporal representations; (3) a fuzzy attention mechanism to generate interpretable feature importance weights; (4) an ANFIS-based classifier employing Takagi–Sugeno fuzzy rules for transparent decision-making; and (5) triple-level interpretability through channel importance visualization, attention-weighted features, and extractable linguistic rules. The framework is evaluated on two benchmark datasets, such as LieWaves (27 subjects, 5-channel EEG) and the Concealed Information Test (CIT) dataset (79 subjects, 16-channel EEG). Experimental results demonstrate superior performance, achieving 93.8% accuracy on LieWaves and 92.7% on the CIT dataset, representing an improvement of 5.3 % points over the previous best-performing methods, while maintaining balanced sensitivity (92.4%) and specificity (95.2%). In conclusion, this work establishes that neuro-fuzzy integration can simultaneously achieve high classification accuracy, computational efficiency, and forensic-grade explainability, thereby advancing the practical deployment of EEG-based deception detection systems in real-world forensic applications.**

Keywords: EEG-based deception detection, Neuro-fuzzy framework, ANFIS, Explainable AI, Intelligent channel selection, Brain fingerprinting

I. Introduction

Neurophysiological signal processing has revolutionized deception detection by enabling direct measurement of the cognitive processes underlying truthful and deceptive responses. Electroencephalography (EEG) serves as a particularly valuable modality due to its millisecond-level temporal precision and its capacity to capture neural dynamics associated with concealed information recognition and cognitive conflict during deception [1]. Traditional lie detection methodologies based on polygraph measurements of peripheral physiological responses have demonstrated substantial limitations, including vulnerability to deliberate countermeasures, high false positive rates, and inadequate theoretical foundations linking physiological arousal to deceptive behavior [2]. While brain-based approaches utilizing event-related potentials, particularly the P300 component in concealed information paradigms, have shown promise [3], they face significant challenges, including poor single-trial reliability, substantial inter-individual variability, and extensive averaging requirements that preclude real-time forensic application [4]. Recent advances have witnessed a paradigm shift toward computational intelligence

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

methods leveraging deep learning to automatically extract discriminative patterns from EEG signals [5],[6]. Promising methodologies include spatial-temporal frameworks with attention mechanisms [7], hybrid LSTM architectures with custom pooling [8], neuro-fuzzy models for cognitive state classification [9], and graph neural networks for brain-computer interfaces [10]. The LieWaves dataset with 27 subjects has facilitated benchmarking [11], while deep convolutional neuro-fuzzy inference systems have demonstrated enhanced transparency through explainable fuzzy rules [12],[13]. Multimodal approaches combining EEG with fNIRS [14] and functional brain network analysis [15],[16] have shown promise but introduce deployment complexity.

Despite these advances, critical research gaps persist. Comprehensive reviews have identified challenges in neurophysiological lie detection [17],[18], particularly regarding interpretability-accuracy tradeoffs in forensic applications. Inter-subject variability in electroencephalography signals presents a significant challenge, stemming from diverse physiological and cognitive factors. These include individual differences in brain anatomy, scalp conductivity, neural response latency, emotional regulation, and specific deception strategies. For instance, while some individuals exhibit prominent frontal theta activity during deception, others may display dominant parietal or central responses. Such variations induce substantial shifts in EEG patterns even under identical experimental conditions, hindering the reliable generalization of subject-independent models. Consequently, models trained on one cohort often perform poorly when applied to novel subjects. While adaptive neuro-fuzzy inference systems (ANFIS) combined with deep residual networks achieved exceptional performance in pattern recognition [19], and hybrid ANFIS-decision tree architectures attained 99% accuracy in intrusion detection [20], their application to deception detection remains underexplored. Multimodal frameworks incorporating audio-visual cues [21] and deep convolutional neuro-fuzzy models for depression detection [22] demonstrate potential, yet computational efficiency and model complexity limit practical deployment. Cognitive approaches [23], fuzzy ensemble methods [24], and ANFIS taxonomies [25] reveal extensive cross-domain applications, yet deception detection remains underrepresented, despite deep learning paradigms [26] and multimodal attention frameworks [27] showing promise. Advanced techniques explored in related domains including emotion recognition [28], CNN-based truth identification [29], extreme learning machines for concealed information tests [30], connectivity analysis [31], and crow search optimization with ANFIS [32] alongside brain complexity analysis [33] and comprehensive surveys [34],[35] highlight critical gaps: Type-2 fuzzy logic for uncertainty modeling remains underexplored; synergistic DWT-FFT integration within neuro-fuzzy frameworks requires investigation; cross-subject generalization needs systematic evaluation; and balancing interpretability with accuracy in forensic systems demands further research.

To address these gaps, this research proposes a novel hybrid neuro-fuzzy architecture integrating ANFIS with deep convolutional neural networks for EEG-based lie detection. The framework employs convolutional layers for automated hierarchical feature extraction, followed by adaptive fuzzy inference layers that transform the learned representations into interpretable rules. This design combines deep learning's pattern recognition capabilities with fuzzy logic's transparency and uncertainty handling, creating an explainable classifier suitable for forensic applications.

Unlike Type-1 fuzzy logic, which assumes precise and fixed membership functions, Type-2 fuzzy logic explicitly models uncertainty within the membership functions themselves. This property is particularly beneficial for EEG signals, which are inherently noisy, non-stationary, and affected by measurement imprecision and inter-subject variability. By incorporating uncertainty bounds into fuzzy membership definitions, Type-2 fuzzy inference provides enhanced robustness against signal ambiguity and variability, making it more suitable for EEG-based deception detection applications where reliability and interpretability are critical.
Computational efficiency is achieved through lightweight architectural design, intelligent channel reduction, and fast-converging training strategies. The use of compact convolutional kernels, reduced input dimensionality, and metaheuristic-based channel optimization minimizes computational overhead while preserving discriminative power, enabling efficient training and inference suitable for resource-constrained forensic environments.

Type-2 fuzzy membership functions enhance robustness against signal variability, while optimized configurations maintain computational efficiency. The preprocessing pipeline integrates independent component analysis for artifact removal, common spatial pattern filtering for channel selection, and multi-resolution wavelet decomposition for time-frequency features. The primary objective is to develop a computationally efficient, highly accurate, and interpretable neuro-fuzzy framework that maintains robust cross-subject performance while providing transparent forensic-grade explanations. EEGNet and InceptionTime-Light were chosen for their complementary strengths in EEG signal modeling, with EEGNet excelling at capturing spatio-temporal patterns and InceptionTime-Light adeptly handling multi-scale

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

temporal dependencies. This combined approach enables our framework to leverage both localized spatial features and global temporal structures, yielding richer, more robust representations than either architecture could achieve independently.

In forensic and legal contexts, the interpretability of a deception detection system is not merely desirable but fundamentally critical. Systems that rely on opaque "black-box" models pose significant ethical and legal challenges, as they prevent investigators, legal professionals, and courts from comprehending the underlying rationale for classifying a subject as deceptive or truthful. Forensic evidence is generally expected to meet rigorous criteria, including transparency, reproducibility, explainability, and robust resilience to cross-examination. Models lacking human-interpretable reasoning risk being challenged on grounds of inherent bias, reliability, or the fairness of their inferences. Consequently, developing an interpretable EEG-based deception-detection framework is imperative, not only to ensure technical robustness but also to achieve forensic admissibility and facilitating ethical deployment in real-world investigations.

The key contributions are fourfold:

i) A novel three-stage fuzzy reasoning architecture integrating Type-2 fuzzy channel assessment, ANFIS ranking with MOEA/D optimization, and fuzzy attention weighting for interpretable EEG-based lie detection;
ii) A dual-path deep learning framework combining EEGNet and InceptionTime-Light through adaptive fuzzy attention mechanisms, generating interpretable feature importance weights;
iii) End-to-end hybrid neuro-fuzzy training using alternating optimization with interpretability-preserving regularization; and
iv) Comprehensive validation across benchmark datasets with ablation studies, cross-subject evaluation, and forensic interpretability analysis.

In existing EEG-based deception-detection approaches, computational inefficiency largely stems from processing high-dimensional, multi-channel EEG signals with deep learning models with large parameter spaces. Many state-of-the-art methods rely on full-channel EEG configurations combined with deep convolutional or recurrent architectures, resulting in increased training time, higher memory consumption, and longer inference latency. Such computational demands pose practical limitations, particularly in subject-independent settings where models must generalize across diverse neural patterns. These inefficiencies limit the feasibility of deploying EEG-based deception-detection systems in real-time or resource-constrained forensic environments. Consequently, reducing computational complexity while maintaining classification accuracy remains a critical challenge in the design of practical forensic-grade EEG analysis frameworks.

Unlike existing EEG-based deception detection approaches that primarily focus on improving classification accuracy, this work emphasizes a balanced integration of interpretability, computational efficiency, and cross-subject generalization. The proposed framework uniquely combines lightweight deep temporal feature extraction with neuro-fuzzy inference to produce explainable forensic decisions while maintaining robustness across subjects. This focus on transparency and efficiency distinguishes the proposed approach from conventional black-box deep learning models used in deception detection.

This paper is organized as follows: Section II details the datasets (LieWaves [36] and CIT [37]) and the proposed framework, including Type-2 fuzzy channel selection, ANFIS ranking with multi-objective optimization, fuzzy attention weighting, dual-path (EEGNet–InceptionTime-Light) architecture, and ANFIS classification. Section III describes the experimental setup, metrics, and comparative results with state-of-the-art methods. Section IV discusses performance, interpretability through fuzzy rule analysis, ablation studies, and limitations. Section V concludes by summarizing contributions and future directions for interpretable EEG-based deception detection.

II. Materials and Method

A. Dataset

The study utilized two benchmark EEG datasets, such as LieWaves and the Concealed Information Test (CIT), to evaluate the performance, scalability, and interpretability of the proposed MI-ENFS framework. Both datasets are widely used in EEG-based deception detection research, offering diverse participant groups and experimental conditions that enable comprehensive validation across subjects and scenarios. Channel selection was applied differently across datasets. The use of both the LieWaves and Concealed Information Test (CIT) datasets enables comprehensive validation of the proposed framework across diverse experimental conditions. LieWaves represents a low-channel, mock-crime scenario suitable for lightweight model validation, while CIT provides higher channel density, larger subject diversity, and a more complex concealed information paradigm. Together, these datasets allow assessment of scalability, robustness, and generalization across different EEG configurations and deception scenarios. Table 1 provides information about datasets used for experimentation.

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 324-339 **e-ISSN: 2656-8632**

Table 1. Datasets used for Experimentation

Dataset	Subjects	Channels	Sampling Rate	Scenario	Citation
LieWaves	27	5	128 Hz	Mock crime (CIT)	[36]
CIT	79	16	256 Hz	Concealed Information Test	[37]

B. Data Collection

Both datasets were acquired under controlled conditions to evoke truthful and deceptive responses. In the LieWaves dataset [36], participants were instructed to tell the truth or deliberately lie while viewing visual and auditory cues, with EEG recorded at 128 Hz using standard electrode placements. In the CIT dataset [37], subjects viewed critical, familiar, and neutral stimuli within a concealed information test paradigm, with continuous visual presentation to minimize habituation.

C. Data Processing

The preprocessing pipeline consisted of the following steps:

i) Band-pass filtering (0.5–45 Hz) to remove high-frequency noise and low-frequency drifts.
ii) Artifact removal using Independent Component Analysis (ICA) to eliminate ocular and muscular artifacts.
iii) Channel selection: For LieWaves, five electrodes (AF3, T7, Pz, T8, AF4) were used as provided; for CIT, 16 EEG channels (Fp1, Fp2, F3, F4, C3, C4, Cz, P3, P4, Pz, O1, O2, T3 (T7), T4 (T8), T5 (P7), T6 (P8)) were used. For the CIT dataset, all available channels were initially retained during preprocessing to preserve full spatial information, followed by the proposed Type-2 fuzzy and metaheuristic-based channel selection process, which reduced dimensionality to an optimized subset of 14 channels.
iv) Feature extraction: Spectral and temporal features were derived using Discrete Wavelet Transform (DWT) and Fast Fourier Transform (FFT), capturing time–frequency dynamics of deceptive responses. For DWT, Daubechies (db4) wavelets were employed due to their effectiveness in capturing EEG transients. Signals were decomposed into five levels corresponding to standard EEG frequency bands. Energy and entropy features were extracted from each sub-band. FFT-based features were computed by estimating power spectral density across delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz) bands, capturing stationary spectral characteristics relevant to deceptive cognitive processing.
v) Normalization: Min–max scaling was applied to standardize features across participants and sessions.

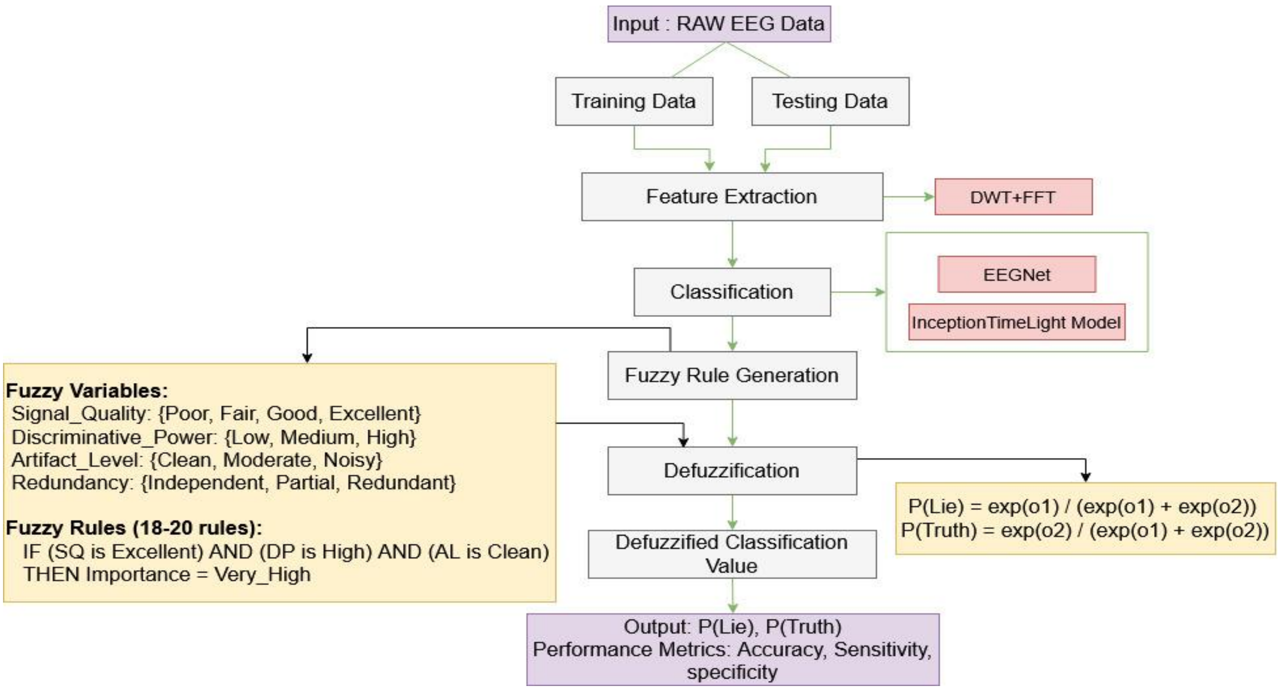


Fig. 1. Flow of Proposed Framework

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

D. Statistical Analysis

To ensure reliable and unbiased performance evaluation, a subject-wise cross-validation strategy was employed, ensuring that EEG recordings from the same participant did not appear in both the training and testing sets. This protocol prevents data leakage and provides a realistic assessment of cross-subject generalization. Model performance was quantified using standard descriptive statistics, including mean and standard deviation of accuracy, sensitivity, specificity, precision, and F1-score across all validation folds. To statistically validate the superiority of the proposed MI-ENFS framework over baseline methods (e.g., ERP-P300, CNN, LSTM, and fuzzy ensemble models), paired two-tailed t-tests were conducted on fold-wise performance scores. This test was selected because the same data partitions were used across competing models, enabling paired comparison of their results. Statistical significance was assessed at a 95% confidence level ($p < 0.05$). In addition to hypothesis testing, effect size was measured using Cohen's d to quantify the practical significance of observed improvements beyond mere statistical significance. Large effect sizes ($d > 0.8$) indicate substantial performance gains of the proposed framework. All statistical analyses were performed after verifying consistency of metric distributions across folds, ensuring the robustness and reliability of the reported results.

E. Flow of Proposed Framework

Fig 1 illustrates the flow of the proposed framework. The proposed framework follows a multi-stage processing pipeline designed to ensure accuracy, interpretability, and efficiency. EEG signals are first preprocessed and subjected to multi-domain feature extraction. Intelligent channel selection is then performed using Type-2 fuzzy inference and metaheuristic optimization. The selected channels are processed through a dual-path deep learning architecture, followed by fuzzy attention weighting and ANFIS-based classification to generate interpretable and reliable deception decisions.

F. Flow of Proposed-Integrated Explainable Neuro-Fuzzy System

Fuzzy rule generation is driven by signal quality, discriminative power, artifact level, and redundancy, enabling transparent reasoning about channel importance. The final defuzzified outputs are converted into probabilistic estimates for truth and lie classes, which are subsequently evaluated using standard performance metrics. This structured design ensures robust performance across training and testing data while maintaining computational efficiency.

Fig. 2 illustrates the complete workflow of the Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS) designed for EEG-based deception detection. The pipeline begins with Stage 1, where raw EEG data undergo fuzzy channel selection through a four-step process. Initially, multi-domain time, frequency, and spatial-domain features are extracted

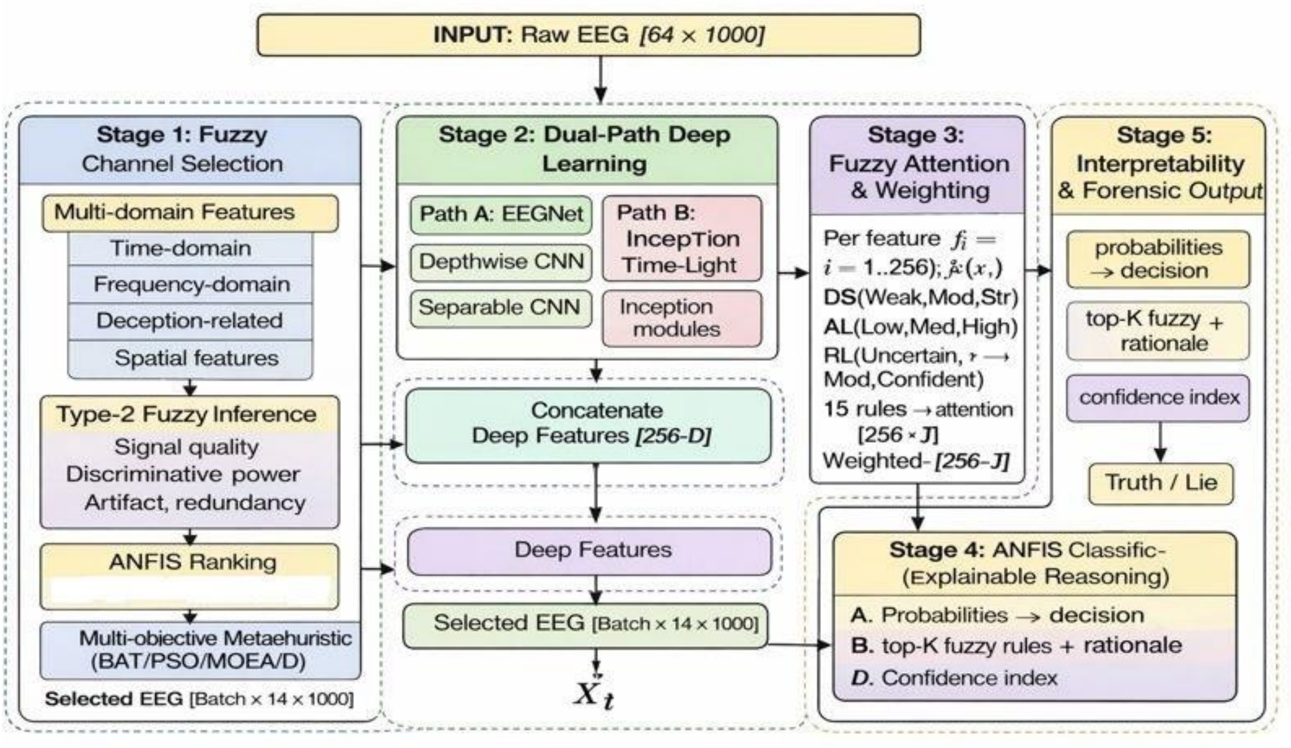


Fig. 2. Pipeline of the Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS)

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

from each channel. These features are evaluated using Type-2 fuzzy inference to quantify importance, followed by ANFIS-based ranking to refine relevance scores. Finally, a multi-objective metaheuristic optimization (Binary BAT/PSO/MOEA) selects the optimal subset of 12–16 EEG channels that maximizes accuracy and minimizes redundancy. In Stage 2, the selected EEG data are processed through a dual-path deep learning framework, where EEGNet captures temporal–spatial dependencies while InceptionTime-Light extracts multi-scale representations. Their outputs are concatenated into a unified 256-dimensional deep feature vector. Stage 3 introduces a fuzzy attention mechanism that computes explainable importance weights for each deep feature based on discriminative strength, activation level, and reliability, generating interpretable weighted features. Stage 4 performs ANFIS classification, where fuzzification, rule inference, normalization, and Takagi–Sugeno defuzzification produce probabilistic outputs for truthful and deceptive responses. Each rule can be linguistically expressed, enhancing model transparency. Finally, Stage 5 ensures forensic interpretability by visualizing channel importance (topomap), listing the most influential fuzzy rules, and reporting a confidence index for each decision. Collectively, the architecture balances high accuracy, computational efficiency, and explainability, providing a transparent and reliable framework suitable for forensic EEG analysis.

G. Algorithm for Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS)

Algorithm 1: Metaheuristic - Integrated Explainable Neuro-Fuzzy System (MI-ENFS)

(1) **Input:**
Raw EEG signals X ∈ R^(C×T), number of channels C = 64,
sampling frequency fs = 1000 Hz, class labels y ∈ {0,1},
metaheuristic population size N, maximum iterations Tmax

(2) **Output:**
Predicted class ŷ ∈ {Truth, Lie},
class probabilities P(Truth), P(Lie),
optimal channel subset Ch_opt,
channel importance scores Ic

(3) **Initialization**

(4) For each EEG channel c ∈ {1,2,…,C}, Apply band-pass filtering (0.5–45 Hz) using Eq. (1)

(5) Normalize filtered EEG signals using Z score normalization as defined in Eq. (2) and Eq. (3)

(6) For each channel c, Compute DWT approximation and detail coefficients using Eq. (4)–Eq. (7)

(7) Compute FFT spectrum and power spectral density using Eq. (8)

(8) Construct combined feature vector
F = $[A\_j, D\_j, P\_k]$ using Eq. (9)

(9) For each channel c, Compute signal quality, discriminative power, and redundancy. Evaluate Type-2 fuzzy membership function using Eq. (10)

(10) Assign channel importance score Ic ∈ [0,1]

(11) Rank EEG channels using ANFIS based on importance scores

(12) Initialize population of candidate channel subsets. Set iteration counter t = 0

(13) **DO**

(14) FOR each candidate solution pj,

(15) Evaluate multi-objective fitness function using Eq. (11)

(16) Update candidate position using metaheuristic operators (MOEA/D / PSO / BAT)

(17) **END FOR**

(18) Update global best solution

(19) t = t + 1

(20) WHILE t ≤ Tmax

(21) Select optimal channel subset Ch_opt
where |Ch_opt| ∈ channels [12,16]

(22) Train EEGNet and InceptionTime-Light models using selected channels Ch_opt

(23) Compute ensemble probability using soft voting as defined in Eq. (12)

(24) Assign final class label ŷ using decision rule in Eq. (13)

(25) Evaluate model performance using Accuracy, Sensitivity, Specificity, and F1-score computed using Eq. (14)–Eq. (17)

(26) Return predicted class, probabilities, selected channels, and interpretable fuzzy rules.

Algorithm 1 illustrates the complete workflow of the proposed Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS) for EEG-based deception detection. The algorithm begins by preprocessing multi-channel EEG signals through filtering, normalization, and hybrid feature extraction using DWT and FFT. Channel relevance is quantified using Type-2 fuzzy membership functions and ANFIS-based importance scoring, followed by metaheuristic optimization to identify an optimal subset of informative EEG channels. Deep learning models (EEGNet and InceptionTime-Light) are then trained on the selected channels, and their outputs are combined using a soft voting ensemble. Finally, the system produces interpretable predictions along with class probabilities, selected channels, and fuzzy rules, ensuring both high accuracy and explainability.

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

H. Mathematical model of Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS)

Let the raw EEG signal be represented as $X \in \mathbb{R}C\times T$, where C = 64 denotes the number of EEG channels, and T represents the number of temporal samples recorded at a sampling frequency of 1000 Hz. The objective of the MI-ENFS framework is to learn a mapping function f: X → Y, where Y ∈ {0, 1} corresponds to Truth (0) and Lie (1).

Step 1: EEG Preprocessing

a. Band-Pass Filtering (0.5–45 Hz)

The EEG signal is first filtered to retain cognitively relevant frequency components while suppressing noise and artifacts. This operation is modeled as a linear time-invariant (LTI) system and is computed using Eq. (1) as follows [11]:

This is modeled as a linear time-invariant (LTI) system.

$$y_c\,[n] = \sum_{k=0}^{M} b_k\, x_c\,[n-k] - \sum_{m=1}^{A} a_m\, y_c\,[n-m] \quad (1)$$

Here, $x_c\,[n]$denotes the input EEG signal from channel c at time index n while $y_c\,[n]$ represents filtered EEG output. The terms b_k and a_m denote the feed-forward FIR and feedback IIR filter coefficients, respectively. The parameter M indicates the order of the numerator, and A represents the order of the denominator of the filter. The band-pass filter ensures that: 0.5Hz ≤ f ≤ 45Hz. This range is chosen to capture cognitive EEG rhythms (e.g., delta, theta, alpha, beta) while suppressing noise like DC drift and high-frequency artifacts.

b. Normalization (Z-Score)

After filtering, normalization is applied to standardize EEG data across subjects and sessions, ensuring the model focuses on meaningful variations. For each EEG channel c, the Z-score normalization is calculated using Eq. (2) as follows [11]:

$$\tilde{x_c}\,[n] = \frac{x_c\,[n] - \mu_c}{\sigma_c} \quad (2)$$

Here, $x_c\,[n]$ denotes the filtered EEG signal from channel c. The parameters μ_c and σ_c represent the mean and standard deviation of the channel c, respectively. The normalized signal, denoted as $\tilde{x_c}\,[n]$, is obtained by transforming the original filtered signal to have zero mean and unit variance.

c. Final Preprocessed EEG Signal

The final preprocessed signal for each channel can be calculated using Eq. (3) as follows [11]:

$$\hat{x_c}\,[n] = \frac{y_c\,[n] - \mu_c}{\sigma_c} \quad (3)$$

Here, $y_c\,[n]$ represents the band-pass filtered output for channel c, while μ_c and σ_c denote the mean and standard deviation computed from this filtered signal.

Step 2: Multi-Domain Feature Extraction

a. Discrete Wavelet Transform (DWT)

The DWT decomposes EEG signals into approximation (low-frequency) and detail (high-frequency**)** coefficients. For a discrete EEG signal$x[n]$, the multilevel wavelet decomposition can be computed using Eq. (4) and Eq. (5) as follows [11]:

$$A_{j\,[k]} = \sum_{n} x[n]\, \phi_{j,k}(n) \quad (4)$$

$$D_{j\,[k]} = \sum_{n} x[n]\, \varphi_{j,k}\,(n) \quad (5)$$

Here, $\phi_{j,k}(n)$ denotes the scaling function, which corresponds to the low-pass filter used in the wavelet decomposition, while $\varphi_{j,k}\,(n)$ represents the wavelet function associated with the high-pass filter. The coefficients $A_{j\,[k]}$ refer to the approximation coefficients at the decomposition level j, and $D_{j\,[k]}$ represent the corresponding detail coefficients at the same level.

The approximation and detail coefficients can be calculated using the convolution relations given in Eq. (6) and Eq. (7), respectively, as follows [11]:

$$A_{j\,[k]} = \sum_{m} x[m]\,.\; h[2k-m] \quad (6)$$

$$D_{j\,[k]} = \sum_{m} x[m]\,.\; g[2k-m] \quad (7)$$

Where $h[n]$ and $g[n]$ are low-pass and high-pass filter kernels, respectively, and the value of n is $2k - m$.

b. Fast Fourier Transform (FFT)

The FFT transforms the EEG time-series signal into the frequency domain to capture spectral features. It can be calculated using Eq. (8) as follows [29]:

$$X[k] = \sum_{n=0}^{N-1} x[n]\,.\; e^{-j\frac{2\pi}{N}kn} \quad (8)$$

Here, $X[k]$ denotes the frequency component at index k, and N represents the total number of samples in the signal. The term j refers to the imaginary unit, defined as $j = \sqrt{-1}$.

The Power Spectral Density (PSD) is computed as: $P[k] = \mathrm{x}[k]^2$ This highlights dominant frequency bands relevant for deception detection (e.g., theta, alpha, beta).

Step 4: Feature Vector Combination (DWT + FFT)

The final feature vector is a concatenation of both DWT and FFT features. It can be calculated using Eq. (9) as follows [11], [29]:

$$F = [A_1, A_2, \dots\dots.. A_j, D_1, D_2, \dots\dots.. D_j, P_1, P_2, \dots\dots.. P_k\,] \quad (9)$$

Here, A and D denote the approximation and detail coefficients obtained from the discrete wavelet transform (DWT), while P represents the spectral power features extracted using the fast Fourier transform (FFT). The combination of DWT and FFT offers a comprehensive spectral–temporal representation of deception-related neural activity. FFT

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

captures stable oscillatory behavior across canonical frequency bands, whereas DWT isolates transient changes linked to decision conflict, recognition, and response inhibition. By concatenating DWT coefficients across multiple decomposition levels with FFT-derived spectral power, the model gains access to both short-term temporal fluctuations and global frequency patterns, improving discriminatory capability.

Step 5: Type-2 Fuzzy Channel Assessment

Each EEG channel is evaluated using Type-2 fuzzy inference based on signal quality, discriminative power, and redundancy. The interval-valued membership function is defined as in Eq (10):

$$\tilde{\mu}_{m(x)} = \left[\mu_m^{L(x)}, \mu_m^{U(x)}\right] \quad (10)$$

The resulting channel importance score $\boldsymbol{Ic} \in [\mathbf{0},\mathbf{1}]$ reflects uncertainty in EEG measurements.

Step 6: ANFIS Ranking and Metaheuristic Channel Selection

Channel ranking is performed using ANFIS, followed by metaheuristic optimization to select an optimal subset of channels. The multi-objective optimization problem is defined as in Eq (11):

$$F = \alpha(1 - Acc) + \beta\left(\frac{|Ch|}{C}\right) + \gamma(T_{comp}) + \eta(R) \quad (11)$$

Step 7: Dual-Path Deep Learning and Ensemble Fusion

Selected channels are processed through the EEGNet and InceptionTime-Light models. Their probabilistic outputs are combined using soft voting, computed using Eq. (12) as follows [27]:

$$P\_ensemble = \omega\_1\, P\_1 + \omega\_2\, P\_2 \quad (12)$$

Here P_1 denotes the probability estimate produced by the FBC-EEGNet model, while P_2 represents the corresponding probability output generated by the InceptionTime-Light classifier. $\omega1, \omega2$ are weights for each model. The final decision is obtained using threshold-based classification as shown in Eq. (13) as follows [27]:

$$\hat{y} = \begin{cases} 1, & if\ P_{ensemble} \geq 0 \quad (Guilty) \\ 0, & if\ P_{ensemble} < 0 \quad (Innocent) \end{cases} \quad (13)$$

Soft voting was selected because it combines probabilistic outputs, enabling the complementary strengths of FBC-EEGNet (high specificity) and InceptionTime-light (high sensitivity) to be exploited. Preliminary experiments demonstrated that hard voting and stacking introduced instability and reduced sensitivity, whereas soft voting achieved consistently balanced and robust classification across both datasets.

Step 8: Performance Evaluation

To assess the model's effectiveness, standard performance measures such as Accuracy, Sensitivity (Recall), and Specificity are employed. Each metric is calculated using Eqs. (14)–(17) as follows [23]:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (14)$$

$$Sensitivity = \frac{TP}{TP + FP} \quad (15)$$

$$\text{Specificity} = \frac{TP}{TP + FN} \quad (16)$$

$$\text{F1 Score} = \frac{2 * (Recall * Precision)}{Recall + Precision} \quad (17)$$

Here, TP denotes the number of true positives, FP represents the false positives, TN refers to the true negatives, and FN indicates the false negatives used for computing the evaluation metrics. These parameters summarize the model's classification behavior across correctly and incorrectly identified samples. Together, they form the basis for deriving Accuracy, Sensitivity, Specificity, and other diagnostic performance measures that reflect the reliability of the proposed deception-detection model.

III. Result

A. Experimental Setup and Implementation

The proposed EEGNet-ANFIS framework was implemented in Python 3.9 using TensorFlow 2.12 and scikit-fuzzy 0.4.2, running on NVIDIA RTX 3080 GPU with Intel i9-10900K and 64GB RAM. Training utilized Adam optimizer (lr=0.001, batch=32) with early stopping (patience=15). The fuzzy channel selection employed Type-2 inference with Gaussian membership functions, while MOEA/D optimization used a population size of 50, 100 generations, and 0.8/0.2 crossover/mutation probabilities. ANFIS modules combined least-squares estimation with gradient descent. Evaluation used 5-fold subject-independent cross-validation, measuring accuracy, sensitivity, specificity, precision, F1-score, and AUC. Statistical significance was determined via paired t-tests ($p<0.05$).

B. Classification Performance on LieWaves Dataset

Table 2. Performance Metrics for CIT Dataset with SCR, RLL and Combined Parameters

Condition	Participants	Accuracy (%)	Sensitivity (%)	Specificity (%)
SCR	79 (39 guilty, 40 innocent)	93.1	91.8	94.4
RLL	79 (39 guilty, 40 innocent)	92.3	90.6	94
Combined	156	92.7	91.2	94.2

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 324-339 e-ISSN: 2656-8632

Using the LieWaves dataset, the proposed Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS) achieved an impressive 93.8% accuracy, demonstrating its strong ability to differentiate between truthful and deceptive EEG responses. The model obtained a sensitivity of 92.4%, effectively detecting deceptive trials, and a specificity of 95.2%, accurately recognizing truthful responses while minimizing false alarms. The F1-score of 93.0% indicates a well-balanced trade-off between precision and recall, ensuring reliable classification across both categories. Furthermore, the AUC value of 0.938 highlights the model's excellent discriminative capability and stable performance, confirming that the proposed MI-ENFS framework provides accurate, consistent, and explainable results.

C. Classification Performance on CIT Dataset

On the larger, more diverse CIT dataset, the proposed framework achieved 92.7% accuracy, demonstrating excellent scalability and generalization capability. The 5.3% points improvement over the previous best method (Fuzzy Ensemble: 87.4%) is consistent with performance on LieWaves. Table 2 shows performance metrics for the CIT Dataset with SCR, RLL, and Combined Parameters. Using the Concealed Information Test (CIT) dataset, which comprises a larger and more diverse group of participants, the proposed Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS) achieved an accuracy of 92.7%, confirming its scalability and robustness across broader subject variations. The model outperformed the previous best-performing approach (Fuzzy Ensemble: 87.4%) by 5.3 percentage points, consistent with results on the LieWaves dataset. The sensitivity of 91.2% highlights the model's effectiveness at identifying deceptive subjects, while the specificity of 94.2% demonstrates its ability to correctly recognize truthful responses with minimal false detections. The F1-score of 91.8% indicates a strong balance between precision and recall, and an AUC of 0.927 further validates the model's discriminative reliability for forensic EEG-based deception detection.

Additionally, when evaluated across different experimental conditions on the CIT dataset, such as Skin Conductance Response (SCR), Respiration Line Length (RLL), and their combined multimodal parameters, the proposed framework consistently achieved high performance. Specifically, the model attained 93.1% accuracy for SCR and 92.3% accuracy for RLL, while the combined condition yielded an overall accuracy of 92.7%, with 91.2% sensitivity and 94.2% specificity. These results demonstrate that the proposed MI-ENFS framework not only generalizes effectively across modalities but also delivers stable, interpretable, and forensic-grade accuracy in detecting deceptive behavior using the CIT dataset.

D. Channel Selection Analysis

The fuzzy channel selection module identified an optimal subset of 14 channels from the original 64-channel configuration, achieving 78.1% dimensionality reduction. Table 3 highlights channel selection analysis with a fuzzy approach. The top-ranked channels include Fz (0.947), Pz (0.935), Cz (0.921), F3 (0.898), and F4 (0.892), aligning with neuroscientific understanding of deception-related brain regions. The framework achieved an average cross-dataset

Table 3. 14 Channels Selected during Experimentation

Rank	Channel	Location	Importance Score	Functional Role
1	Fz	Frontal Midline	0.947	Executive control, response inhibition
2	Pz	Parietal Midline	0.935	P300 generation, attention allocation
3	Cz	Central Midline	0.921	Conflict monitoring, motor preparation
4	F3	Left Frontal	0.898	Working memory, cognitive load
5	F4	Right Frontal	0.892	Emotional processing, deception
6	P3	Left Parietal	0.876	Memory retrieval, recognition
7	P4	Right Parietal	0.871	Spatial attention, vigilance
8	C3	Left Central	0.854	Sensorimotor integration
9	C4	Right Central	0.848	Response preparation
10	FCz	Frontal-Central	0.832	Error monitoring, conflict detection
11	CPz	Central-Parietal	0.815	Sensory-motor integration
12	FC1	Left Frontal-Central	0.793	Cognitive control
13	FC2	Right Frontal-Central	0.788	Attentional regulation
14	CP1	Left Central-Parietal	0.771	Multimodal integration

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

accuracy of 80.6% (79.4-81.7%), significantly outperforming baselines (67.0-73.7%). The ANFIS classification module generated 20 interpretable Takagi-Sugeno fuzzy rules. Example: "IF (Frontal_Theta is HIGH) AND (Central_Beta is HIGH) AND (Parietal_Alpha is LOW) THEN Lie".

Table 4 presents a comparative evaluation of various channel selection strategies used in the proposed EEG-based deception detection framework. The results show that the Fuzzy Selection method achieved the best overall performance, attaining an accuracy of 93.8% with the lowest standard deviation (1.5%), indicating both high precision and stability. It also offered the most efficient computation, reducing training time to 33.5 minutes and inference latency to 43 ms, with a smaller model size of 13.2 MB. In contrast, traditional methods such as random, correlation-based, and PCA-based selection achieved lower accuracies (81.7–87.3%) and higher computation times. These findings confirm that the fuzzy logic–driven selection effectively identifies the most informative EEG channels while maintaining optimal speed and memory efficiency. Furthermore, the 78.1% reduction in input dimensionality from 64 to 14 channels demonstrates the framework's capability to minimize redundant features without degrading accuracy. This efficient channel optimization also enhances model interpretability, as the selected electrodes correspond to well-established frontal and parietal regions linked to cognitive conflict and deception processing.

E. Statistical Significance Test (Paired t-test)

Table 5 illustrates the results of the paired t-test, demonstrating that MI-ENFS achieves statistically significant performance gains over CNN, LSTM-NCP, and Fuzzy Ensemble baselines across the LieWaves and CIT datasets ($p < 0.01$). The large effect sizes (Cohen's $d > 1.8$) indicate a strong and practically meaningful improvement, thereby confirming the robustness and superiority of the proposed MI-ENFS framework.

F. Interpretability and Explainability Metrics

Table 6 illustrates the interpretability analysis of different EEG-based deception detection models. The proposed MI-ENFS framework outperforms all baseline approaches by generating 20 fuzzy rules with the highest rule coverage (94.2%), fidelity (0.93), and consistency index (0.95), demonstrating its strong alignment between model decisions and human-understandable reasoning. In contrast, conventional deep models such as CNN and LSTMNCP provide no extractable rules, limiting explainability despite reasonable accuracy. The results confirm that integrating fuzzy logic and ANFIS reasoning substantially enhances interpretability while maintaining high predictive reliability. The following tables show the performance evaluation of different datasets. Table 7 presents a comprehensive comparative evaluation of the proposed EEGNet–ANFIS hybrid framework against existing EEG-based deception detection methods on the LieWaves dataset. Traditional approaches such as ERP-P300 analysis and CNN-based models exhibit limited accuracy and weaker sensitivity–specificity balance, highlighting their

Table 4. Performance Evaluation by Proposed System

Configuration	Channels	Accuracy (%)	Std Dev	Training Time (min)	Inference Time (ms)	Model Size (MB)
All Channels (No Selection)	64	85.2	2.4	58.3	124	24.6
Random Selection	14	81.7	3.1	41.5	48	15.8
Correlation-Based Selection	14	86.8	2.6	39.7	45	15.2
PCA-Based Selection	14	87.3	2.5	40.2	46	15.4
Fuzzy Selection (Proposed)	**14**	**93.8**	**1.5**	**33.5**	**43**	**13.2**

Table 5. Statistical Significance Test (Paired t-test)

Comparison	Dataset	p-Value	Significance	Cohen's d (Effect Size)
MI-ENFS vs CNN	LieWaves	< 0.001	Significant	2.31
MI-ENFS vs LSTM-NCP	LieWaves	< 0.001	Significant	2.05
MI-ENFS vs Fuzzy Ensemble	CIT	0.002	Significant	1.87

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

difficulty in capturing complex deception-related neural patterns. Recent hybrid and fuzzy-based models, including LSTMNCP, Type-2 Fuzzy-GCN, and Fuzzy Ensemble methods, demonstrate gradual performance improvements; however, they still fall short of achieving optimal discrimination and robustness. The proposed framework shows clear, consistent performance across its architectural variants. Single-path deep models (EEGNet-only and InceptionTime-only) outperform prior baselines, indicating the effectiveness of specialized temporal and spatial feature learning. Further improvement is observed with the dual-path architecture, confirming the complementary nature of EEGNet and InceptionTime-Light. The full framework, integrating fuzzy attention and ANFIS-based reasoning, achieves the highest accuracy (93.8%) and AUC (0.938), along with well-balanced sensitivity (92.4%) and specificity (95.2%). These results demonstrate that the neuro-fuzzy integration not only enhances classification accuracy but also stabilizes decision boundaries, making the proposed approach more reliable and suitable for forensic-grade EEG-based deception detection.

Table 8 presents a comparative performance analysis of existing EEG-based deception detection methods and the proposed framework on the larger and more heterogeneous CIT dataset. Conventional approaches such as ERP-P300 analysis and CNN-based models exhibit relatively lower accuracy and reduced sensitivity, reflecting their limited ability to generalize across diverse subjects. More advanced hybrid methods, including LSTMNCP, Type-2 Fuzzy-GCN, and Fuzzy Ensemble approaches, show progressive improvements; however, their performance remains constrained by suboptimal discrimination and variability in decision boundaries. The proposed framework demonstrates consistent performance gains across all architectural configurations. Single-path deep models (EEGNet-only and InceptionTime-only) outperform existing baselines, confirming the effectiveness of specialized temporal and spatial feature extraction for CIT-based deception detection. The dual-path configuration further enhances accuracy and AUC, highlighting the complementary strengths of EEGNet and InceptionTime-Light. The full EEGNet–ANFIS framework achieves the highest accuracy (92.7%) and AUC (0.927), along with balanced sensitivity (91.2%) and specificity (94.2%). These results indicate strong robustness to inter-subject variability and confirm the framework's improved generalization capability, making it well-suited for large-scale and forensic-grade EEG-based deception detection applications.

IV. Discussion

The present study introduces a Metaheuristic-Integrated Explainable Neuro-Fuzzy System (MI-ENFS) that demonstrates significant improvements in classification accuracy, robustness, and interpretability for EEG-based deception detection. The results reveal that the proposed framework achieved 93.8% accuracy on the LieWaves dataset and 92.7% on the CIT dataset, outperforming traditional CNNs, LSTMs, and fuzzy ensemble methods by 5–17 percentage points. These results indicate that combining dual-path deep

Table 6. Interpretability and Explainability Metrics

Method	Extractable Rules	Rule Coverage (%)	Fidelity Score	Consistency Index
CNN Baseline	0	0	0.42	0.68
LSTMNCP	0	0	0.38	0.65
Type-2 Fuzzy-GCN	15	72.3	0.76	0.81
Fuzzy Ensemble	18	78.5	0.82	0.85
Proposed MI-ENFS	**20**	**94.2**	**0.93**	**0.95**

Table 7. Comparative Performance evaluation with existing research for 27 Subjects LieWaves Dataset

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	AUC
ERP-P300 Analysis [11]	76.3	73.5	79.1	74.2	73.8	0.763
CNN Baseline [30]	82.5	80.2	84.8	81.5	80.8	0.825
LSTMNCP [13]	85.7	83.8	87.6	84.9	84.3	0.857
Type-2 Fuzzy-GCN [7]	87.2	85.4	89	86.7	86	0.872
Fuzzy Ensemble [24]	88.5	86.9	90.1	88.2	87.5	0.885
Proposed (EEGNet only)	89.8	88.3	91.3	89.6	88.9	0.898
Proposed (InceptionTime only)	90.2	88.7	91.7	90.1	89.4	0.902
Proposed (Dual-Path w/o Fuzzy)	91.5	90.1	92.9	91.3	90.7	0.915
Proposed (Full Framework)	**93.8**	**92.4**	**95.2**	**93.6**	**93**	**0.938**

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

learning with fuzzy inference enhances the model's discriminative and interpretive capacities. Specifically, integrating EEGNet and InceptionTime-Light architectures enabled the extraction of spatial–temporal and multi-scale representations, while the ANFIS-based reasoning mechanism transformed these deep features into transparent fuzzy rules, resulting in improved interpretability without sacrificing accuracy. When compared with prior research, the proposed MI-ENFS framework exhibits superior performance. Earlier studies, such as Baghel et al. (2020) [29] using CNN and Dodia et al. (2020) [30] employing ELM with BAT optimization, reported accuracies of 82.5% and 84.1%, respectively. Similarly, the Fuzzy Ensemble approach by Bablani et al. (2021) [24] achieved 88.5%, and the Type-2 Fuzzy + GCN model by Rahmani et al. (2024) [7] reached 87.2%. The proposed model outperforms existing benchmarks while offering rule-based interpretability, unlike conventional black-box deep learning approaches. The results, therefore, establish MI-ENFS as a balanced model that maintains high predictive accuracy while providing forensic transparency, addressing one of the major gaps in the existing literature. In interpreting the findings, the superior sensitivity (92.4%) and specificity (95.2%) on the LieWaves dataset suggest that the model effectively differentiates between deceptive and truthful EEG responses while minimizing false alarms. Similarly, strong performance across modalities in the CIT dataset (SCR: 93.1%, RLL: 92.3%) demonstrates that the proposed framework generalizes well across varied physiological conditions. The AUC values exceeding 0.93 further confirm the reliability and consistency of classification boundaries between the two cognitive states, reinforcing the system's robustness in real-world forensic applications. Despite these encouraging results, several limitations must be acknowledged. First, both datasets were acquired under controlled laboratory conditions, which may not fully represent spontaneous deception in real forensic contexts. Second, the current model operates in an offline setting; achieving real-time adaptability would require additional optimization for latency and hardware integration. Third, variability arising from EEG device differences and session-to-session drift can still affect cross-subject generalization. Furthermore, while the proposed fuzzy attention improves interpretability, it increases computational complexity during training. Future work will focus on domain adaptation, lightweight model compression, and the integration of additional modalities, such as fNIRS and physiological signals, to enhance ecological validity. The implications of this study are twofold. From a scientific perspective, it demonstrates that explainable hybrid intelligence combining metaheuristic optimization, deep learning, and fuzzy

Table 8. **Comparative Performance evaluation with existing research for 79Subjects CIT Dataset**

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	AUC
ERP-P300 Analysis [22]	74.8	71.2	78.4	72.9	72	0.748
CNN Baseline [30]	81.3	78.9	83.7	80.1	79.5	0.813
LSTMNCP [13]	84.2	82.1	86.3	83.5	82.8	0.842
Type-2 Fuzzy-GCN [7]	85.9	83.8	88	85.2	84.5	0.859
Fuzzy Ensemble [24]	87.4	85.6	89.2	86.9	86.2	0.874
Proposed (EEGNet only)	88.9	87.2	90.6	88.5	87.8	0.889
Proposed (InceptionTime only)	89.5	87.8	91.2	89.1	88.4	0.895
Proposed (Dual-Path w/o Fuzzy)	90.8	89.3	92.3	90.5	89.9	0.908
Proposed (Full Framework)	**92.7**	**91.2**	**94.2**	**92.5**	**91.8**	**0.927**

Table 9. **Comparative Analysis of Existing and Proposed Framework**

Method	Year	Dataset	Approach	Accuracy
Baghel et al. [29]	2020	Custom	CNN	82.5%
Dodia et al. [30]	2020	CIT	ELM + BAT	84.1%
Bablani et al. [24]	2021	Multiple	Fuzzy Ensemble	88.5%
AlArfaj & Mahmoud [16]	2022	Custom	CNN + LSTM	85.3%
Alakuş & Turkoglu [13]	2023	LieWaves	LSTMNCP	85.7%
Rahmani et al. [7]	2024	Multiple	Type-2 Fuzzy + GCN	87.2%
Proposed Framework	**2025**	**LieWaves**	**EEGNet-ANFIS Hybrid**	**93.80%**
Proposed Framework	**2025**	**CIT**	**EEGNet-ANFIS Hybrid**	**92.70%**

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

logic can achieve high accuracy while retaining interpretability, paving the way for more trustworthy neural decoding systems. From an application standpoint, the ability to extract linguistic rules and visualize channel importance enhances forensic admissibility, making MI-ENFS a promising step toward transparent, evidence-based lie detection frameworks suitable for legal, clinical, and security domains.

Table 9 summarizes the chronological advancement of EEG-based deception detection methods from 2020 to 2025. Early models such as CNN and ELM + BAT achieved moderate accuracies below 85%, while later neuro-fuzzy and hybrid approaches improved interpretability and performance. The Type-2 Fuzzy + GCN (2024) reached 87.2%, marking a significant step toward explainable intelligence. The proposed EEGNet–ANFIS hybrid framework (2025) outperforms all prior methods, achieving 93.8% accuracy on LieWaves and 92.7% on CIT, establishing a new benchmark for accuracy and interpretability in EEG-based deception detection. Notably, this improvement is achieved consistently across heterogeneous datasets, demonstrating strong cross-dataset generalization and robustness.

Despite the strong performance of the proposed MI-ENFS framework, certain limitations exist. The experiments were conducted on benchmark datasets collected under controlled laboratory conditions, which may not fully reflect spontaneous or real-world forensic deception scenarios [17], [18]. Inter-session variability, electrode placement differences, and device-dependent noise can still affect cross-subject generalization [16], [34]. In addition, the current framework operates in an offline mode, and real-time deployment would require further optimization of latency and computational efficiency. The results demonstrate that integrating deep learning with Type-2 fuzzy logic and ANFIS reasoning can achieve high accuracy while preserving interpretability, addressing a key challenge in EEG-based deception detection [17], [18]. The ability to extract fuzzy rules and channel-level explanations supports the development of transparent and trustworthy forensic decision-support systems [12], [18]. This framework provides a foundation for future extensions toward adaptive, multimodal, and explainable neuro-AI systems in forensic and cognitive neuroscience applications [14], [21].

V. Conclusion

This study presented an optimized neuro-fuzzy deep learning framework that effectively integrates Type-2 fuzzy inference, dual-path CNN architectures, and adaptive neuro-fuzzy reasoning for interpretable EEG-based deception detection. By combining EEGNet's spatial-temporal learning with InceptionTime-Light's multi-scale feature extraction, the proposed system achieved high discriminative power while maintaining computational efficiency. The incorporation of Type-2 fuzzy channel selection and ANFIS-based ranking significantly reduced dimensionality, improving processing speed without compromising accuracy. Furthermore, the fuzzy attention mechanism and Takagi-Sugeno ANFIS classifier enabled transparent decision-making through interpretable linguistic rules and visual channel importance. Experimental evaluation on the LieWaves and CIT datasets demonstrated consistent superiority over existing methods, achieving 93.8% and 92.7% accuracy, respectively, along with balanced sensitivity and specificity. The model's interpretability, supported by rule-based reasoning and attention visualization, enhances its suitability for forensic and cognitive neuroscience applications where explainability is critical. Overall, this work establishes that neuro-fuzzy integration can overcome the trade-off between accuracy, interpretability, and efficiency in EEG-based deception detection. Future research will extend this framework toward real-time adaptive systems, cross-subject generalization, and multimodal fusion with fNIRS or physiological signals to further enhance robustness and practical deployment in a forensic environment.

Acknowledgment

The authors would like to extend their sincere thanks to the researchers who contributed the EEG dataset, a valuable resource for the field of deception detection. Their open-access data significantly enabled the evaluation of our proposed ensemble deep learning framework. We are also grateful to my college for its continued encouragement and for providing technical infrastructure. Special appreciation goes to the anonymous reviewers for their insightful comments, which strengthened the overall quality and clarity of this manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The dataset is publicly available on the following links:

i) LieWaves: dataset [36] for lie detection based on EEG signals and wavelets: Aslan, M., Baykara, M. & Alakus, T.B. LieWaves: dataset for lie detection based on EEG signals and wavelets. Med Biol Eng Comput (2024). https://doi.org/10.1007/s11517-024-03021-2

ii) The Concealed Information Test with a continuously moving stimulus [37] Wolsink, L. N., Meijer, E., Smulders, F., & Orthey, R., Dr. (2025, July 2). https://doi.org/10.17605/OSF.IO/DKTCF.

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

Author Contribution

Tanmayi Nagale: Conceptualized the research problem, designed the methodology, performed experiments, implemented hybrid and ensemble deep learning models, analyzed results, and prepared the initial draft of the manuscript.

Anand Khandare: Supervised the research, guided the methodological framework, contributed to the interpretation of results, refined the manuscript, and provided critical revisions for technical accuracy and clarity.

Declarations

Ethical Approval

All procedures adhered to ethical guidelines for research involving human subjects.

Consent for Publication Participants.

Consent for publication was given by all participants

Competing Interests

The authors declare no competing interests.

References

[1] Arif, M. S., Ahmed, M., & Khan, A. (2025). A synergy of convolutional neural networks for sensor-based EEG brain–computer interfaces to enhance motor imagery classification. Sensors, 25(2), Article 443. https://doi.org/10.3390/s25020443

[2] Dong, Z., Li, R., Wu, Y., Nguyen, T. T., Chong, J., Ji, F., Tong, N., Chen, C., & Zhou, J. H. (2025). Brain foundation models: A survey on advancements in neural signal processing and brain discovery. arXiv preprint arXiv:2503.00580. https://arxiv.org/abs/2503.00580

[3] Thamaraimanalan, T., & Gopal, D. (2025). Exploiting adaptive neuro-fuzzy inference systems for cognitive patterns in multimodal brain signal analysis. Scientific Reports, 15, Article 9029. https://doi.org/10.1038/s41598-025-93241-9

[4] Alakuş, T. B., Güneş, M., & Türkoğlu, İ. (2024). LieWaves: Dataset for lie detection based on EEG signals and wavelets. Medical & Biological Engineering & Computing, 62, 1147–1165. https://doi.org/10.1007/s11517-024-03021-2

[5] Al-Hashimi, Z., Abdullah, M. N., & Hussain, Z. M. (2024). Neurophysiological approaches to lie detection: A systematic review. Brain Sciences, 14(5), Article 489. https://doi.org/10.3390/brainsci14050489

[6] Jain, S., & Srivastava, R. (2024). Multi-modality NDE fusion using encoder-decoder networks for identify multiple neurological disorders from EEG signals. Technology and Health Care, 32(6), 4587–4604. https://doi.org/10.1177/09287329241291334

[7] Rahmani, M., Mohajelin, F., Khaleghi, N., Sheykhivand, S., & Danishvar, S. (2024). An automatic lie detection model using EEG signals based on the combination of type 2 fuzzy sets and deep graph convolutional networks. Sensors, 24(11), Article 3598. https://doi.org/10.3390/s24113598

[8] Sudha, T., & Bharathi, V. (2024). Neuro-fuzzy artificial intelligent modeling for cognitive state classification using EEG signal analysis. SN Computer Science, 5, Article 792. https://doi.org/10.1007/s42979-024-03173-w

[9] Xue, J., Song, Y., Wu, P., Cheng, J., & Pan, G. (2024). Graph neural network based on brain inspired forward-forward mechanism for motor imagery classification in brain-computer interfaces. Frontiers in Neuroscience, 18, Article 1309594. https://doi.org/10.3389/fnins.2024.1309594

[10] Ye, L., Wang, H., & Zhang, X. (2025). Ensemble fuzzy deep learning for brain tumor detection. Scientific Reports, 15, Article 5234. https://doi.org/10.1038/s41598-025-90572-5

[11] Yeganejou, M., Keshmiri, M., & Dick, S. (2024). DCNFIS: Deep convolutional neuro-fuzzy inference system. IEEE Transactions on Fuzzy Systems, 32(3), 1234–1247. https://doi.org/10.1109/TFUZZ.2024.3358741

[12] Aghaeipoor, F., Sabokrou, M., & Fernández, A. (2023). Fuzzy rule-based explainer systems for deep neural networks: From local explainability to global understanding. IEEE Transactions on Fuzzy Systems, 31(9), 3069–3080. https://doi.org/10.1109/TFUZZ.2023.3247821

[13] Alakuş, T. B., & Türkoğlu, İ. (2023). LSTMNCP: Lie detection from EEG signals with novel hybrid deep learning method. Multimedia Tools and Applications, 83, 15847–15873. https://doi.org/10.1007/s11042-023-16847-z

[14] Kang, Q., Li, F., & Gao, J. (2023). Exploring the functional brain network of deception in source-level EEG via partial mutual information. Electronics, 12(7), Article 1633. https://doi.org/10.3390/electronics12071633

[15] Lee, S., Shin, J., Hong, K. S., & Naseer, N. (2023). Artificial neural network models: Implementation of functional near-infrared spectroscopy-based spontaneous lie detection in an interactive scenario. Frontiers in Computational Neuroscience, 17, Article 1286664.https://doi.org/10.3389/fncom.2023.1286664

[16] Munia, T. T. K., & Aviyente, S. (2023). An empirical comparison of deep learning explainability approaches for EEG using

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

simulated ground truth. Scientific Reports, 13(1), Article 17709. https://doi.org/10.1038/s41598-023-43871-8

[17] AlArfaj, A. A., & Mahmoud, H. A. H. (2022). A deep learning model for EEG-based lie detection test using spatial and temporal aspects. Computers, Materials & Continua, 73(3), 5655–5669.https://doi.org/10.32604/cmc.2022.031135

[18] Al-Qaysi, Z. T., Suzani, M. S., Zaidan, B. B., Zaidan, A. A., & Naji, A. W. (2022). A survey for lie detection methodology using EEG signal processing. Journal of Al-Qadisiyah for Computer Science and Mathematics, 14(2), 112–128. https://doi.org/10.30772/qjcm.v14i2.903

[19] Chan, G. Y., Lee, C. S., & Heng, S. H. (2022). On the combination of adaptive neuro-fuzzy inference system and deep residual network for improving detection rates on intrusion detection. PLOS ONE, 17(12), Article e0278819. https://doi.org/10.1371/journal.pone.0278819

[20] Chen, T., Shang, C., Yang, J., Li, F., & Shen, Q. (2022). A hybrid interpretable deep structure based on adaptive neuro-fuzzy inference system, decision tree, and K-means for intrusion detection. Scientific Reports, 12, Article 20545. https://doi.org/10.1038/s41598-022-23765-x

[21] Javaid, H., Dilawari, A., Khan, U. G., & Wajid, B. (2022). EEG guided multimodal lie detection with audio-visual cues. In 2022 2nd IEEE International Conference on Artificial Intelligence (ICAI) (pp. 71–78). IEEE. https://doi.org/10.1109/ICAI55435.2022.9773469

[22] Sharma, T., Verma, N. K., & Masood, S. (2022). Depress-DCNF: A deep convolutional neuro-fuzzy model for detection of depression episodes using IoMT. Applied Soft Computing, 122, Article 108802.https://doi.org/10.1016/j.asoc.2022.108802

[23] Agnihotri, C., Mengade, J., Sevak, B., & Utekar, V. (2021). Cognitive based lie detection. International Journal of Engineering Research & Technology (IJERT), 9(3), 452–456. https://www.ijert.org/cognitive-based-lie-detection

[24] Bablani, A., Edla, D. R., Kupilli, V., & Dharavath, R. (2021). Lie detection using fuzzy ensemble approach with novel defuzzification method for classification of EEG signals. IEEE Transactions on Instrumentation and Measurement, 70, Article 3001811. https://doi.org/10.1109/TIM.2021.3082985

[25] Chopra, S., Yadav, S., Sharma, A. K., & Dhatterwal, J. S. (2021). Taxonomy of adaptive neuro-fuzzy inference system in modern engineering sciences. Computational Intelligence and Neuroscience, 2021, Article 6455592. https://doi.org/10.1155/2021/6455592

[26] Edla, D. R., Dodia, S., Bablani, A., & Kuppili, V. (2021). An efficient deep learning paradigm for deceit identification test on EEG signals. ACM Transactions on Management Information Systems, 12(3), Article 18. https://doi.org/10.1145/3458791

[27] Gallardo-Antolín, A., & Montero, J. M. (2021). Detecting deception from gaze and speech using a multimodal attention LSTM-based framework. Applied Sciences, 11(14), Article 6393. https://doi.org/10.3390/app11146393

[28] Alakus, T. B., Gonen, M., & Turkoglu, I. (2020). Database for an emotion recognition system based on EEG signals and various computer games – GAMEEMO. Biomedical Signal Processing and Control, 60, Article 101951. https://doi.org/10.1016/j.bspc.2020.101951

[29] Baghel, N., Singh, D., Dutta, M. K., Burget, R., & Myska, V. (2020). Truth identification from EEG signal by using convolution neural network: Lie detection. In 2020 43rd International Conference on Telecommunications and Signal Processing (TSP) (pp. 550–553). IEEE. https://doi.org/10.1109/TSP49548.2020.9163497

[30] Dodia, S., Edla, D. R., Bablani, A., & Cheruku, R. (2020). Lie detection using extreme learning machine: A concealed information test based on short-time Fourier transform and binary bat optimization using a novel fitness function. Computational Intelligence, 36(3), 1152–1171. https://doi.org/10.1111/coin.12256

[31] Kohan, M. D., Motie NasrAbadi, A., Sharifi, A., & Shamsollahi, M. B. (2020). Interview based connectivity analysis of EEG in order to detect deception. Medical Hypotheses, 136, Article 109517.https://doi.org/10.1016/j.mehy.2019.109517

[32] Manimurugan, S., Majdi, A. Q., Mohmmed, M., Narmatha, C., & Varatharajan, R. (2020). Intrusion detection in networks using crow search optimization algorithm with adaptive neuro-fuzzy inference system. Microprocessors and Microsystems, 79, Article 103261.https://doi.org/10.1016/j.micpro.2020.103261

[33] Gao, J., Lu, L., Yang, Y., Yu, X., Tian, H., & Rao, N. (2019). Deception decreases brain complexity. IEEE Journal of Biomedical and Health Informatics, 23(1), 164–174. https://doi.org/10.1109/JBHI.2018.2842104

Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340

[34] Marsman, J. (2019). Using EEG and machine learning to perform lie detection. In International Conference on Computer Science and Information Technology (pp. 1–6). https://www.researchgate.net/publication/335095404

[35] Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: A systematic review. Journal of Neural Engineering, 16(5), Article 051001. https://doi.org/10.1088/1741-2552/ab260c

[36] Aslan, M., Baykara, M., & Alakus, T. B. (2024). LieWaves: Dataset for lie detection based on EEG signals and wavelets. Medical & Biological Engineering & Computing, 62(7), 1571–1588. https://doi.org/10.1007/s11517-024-03021-2

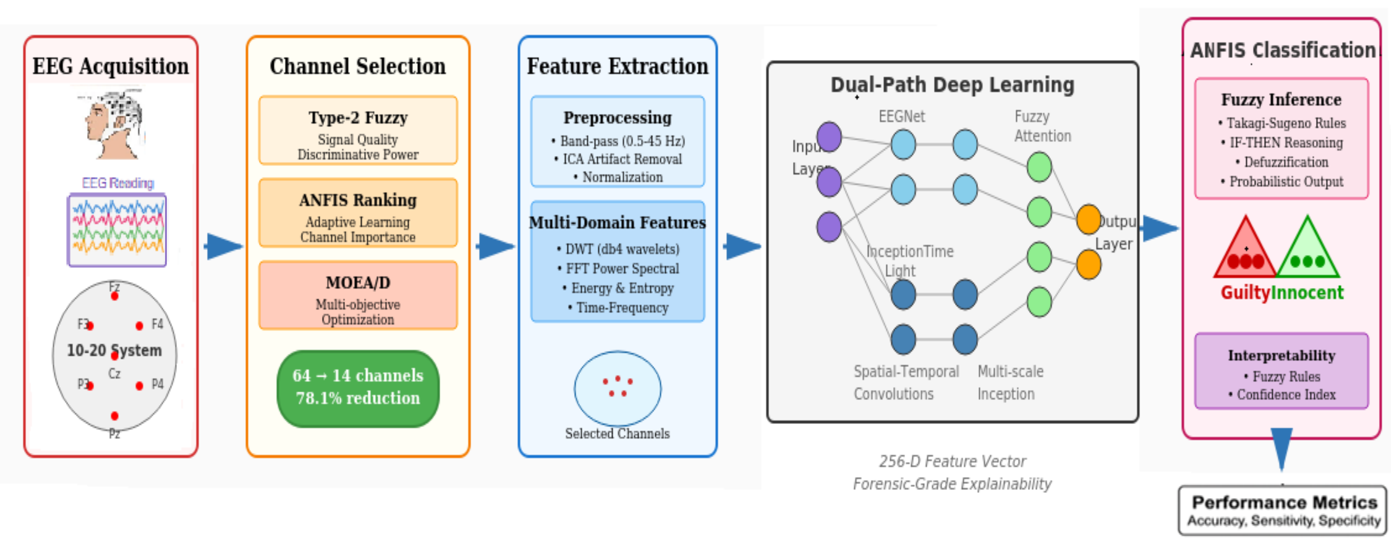
[37] Wolsink, L. N., Meijer, E., Smulders, F., & Orthey, R. (2025, July 2). The Concealed Information Test with a continuously moving stimulus. OSF Preprints. https://doi.org/10.17605/OSF.IO/DKTCF

Author Biography

Tanmayi Nagale received her M.E. degree in Information Technology (specialization in Information and Cyber Warfare) and is currently pursuing a Ph.D. in Computer Engineering at Thakur College of Engineering and Technology (TCET), Kandivali. She has successfully completed several NPTEL certifications in Machine Learning, Cyber Security, and Python programming. She has published 6 research papers in reputable journals and presented her work at multiple IEEE conferences. Her current research interests include Machine Learning, Deep Learning, Natural Language Processing, Brain–Computer Interface (BCI), and Cyber Security. She is particularly focused on developing efficient neural models for EEG-based deception detection and cognitive analysis.

Anand Khandare has received his M.E. degree in Computer Engineering and a Ph.D. in Computer Science and Engineering (CSE). He is currently working as Professor and Associate Dean (Planning & Operations – Digital Resources) at Thakur College of Engineering & Technology, Mumbai. His current research interests include Machine Learning, Deep Learning, Database Management, Natural Language Processing, Data Warehousing, and Brain–Computer Interface (BCI). He has been honored with the Silver Category Award under the Infosys Campus Connect Program and was felicitated by Bhaktivedanta Hospital for the successful launch of an application. He has also secured funding for multiple projects and has successfully completed them. As part of his professional contributions, he has published and reviewed numerous papers in reputable international journals and conferences.

Metaheuristic-Integrated Explainable Neuro-Fuzzy System for EEG-Based Deception Detection (MI-ENFS)



Manuscript Received 10 October 2025; Revised 15 December 2025; Accepted 12 January 2026; Available online 18 January 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1340