Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 152-167 e-ISSN: 2656-8632

RESEARCH ARTICLE

OPEN ACCESS

# CIT-LieDetect: A Robust Deep Learning Framework for EEG-Based Deception Detection Using Concealed Information Test

**Tanmayi Nagale**[1], **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:** **Deception detection with electroencephalography (EEG) is still an open problem as a result of inter-individual variability of brain activity and neural dynamics of deceitful responses. Traditional methods fail to perform well in terms of consistent generalization, and as a result, research has shifted towards exploring sophisticated deep learning methods for Concealed Information Tests (CIT). The objective of the present study is to categorize subjects as guilty or innocent based on EEG measurements and rigorously test model performance in terms of accuracy, sensitivity, and specificity. To achieve this, experiments were conducted on two EEG datasets: the LieWaves dataset, consisting of 27 subjects recorded with five channels (AF3, T7, Pz, T8, AF4), and the CIT dataset, comprising 79 subjects recorded with 16 channels (Fp1, Fp2, F3, F4, C3, C4, Cz, P3, P4, Pz, O1, O2, T3/T7, T4/T8, T5/P7, T6/P8). Preprocessing involved a band-pass filter for noise reduction, followed by feature extraction using the Discrete Wavelet Transform (DWT) and the Fast Fourier Transform (FFT). Three models were evaluated: FBC-EEGNet, InceptionTime-light, and their ensemble. Results indicate that InceptionTime-light achieved the highest accuracy of 79.2% on the CIT dataset, surpassing FBC-EEGNet (70.8%). On the LieWaves dataset, FBC-EEGNet achieved superior performance, with 71.6% accuracy, compared with InceptionTime-light (65.93%). In terms of specificity, FBC-EEGNet reached 93.7% on the CIT dataset, while InceptionTime-light demonstrated balanced performance with 62.5% sensitivity and 87.5% specificity. Notably, the ensemble model provided stable and generalizable outcomes, yielding 70.8% accuracy, 62.5% sensitivity, and 75% specificity on the CIT dataset, confirming its robustness across subject groups. In conclusion, FBC-EEGNet is effective for maximizing specificity, InceptionTime-light achieves higher accuracy, and the ensemble model delivers a balanced trade-off. The implications of this work are to advance reliable EEG-based deception detection and to set the stage for future research on explainable and interpretable models, validated on larger and more diverse datasets.**

**Keywords:** **EEG, Deception Detection, FBC-EEGNet, InceptionTime-light, Ensemble Model.**

RETRACTED

## I. Introduction

Electroencephalography (EEG) has emerged as a powerful tool for studying brain activity due to its non-invasive nature and high temporal resolution. It has been widely applied in domains such as emotion recognition and deception detection. While emotion recognition leverages EEG signals to analyze affective states, deception detection primarily focuses on identifying guilty or innocent responses in Concealed Information Tests (CIT) and related paradigms. Traditional approaches relied on event-related potentials (ERP) such as the P300, which provided valuable neural markers of recognition. However, these methods often struggled to generalize across individuals and required controlled laboratory settings, thereby limiting their real-world applicability. Despite progress, current EEG-based deception-detection techniques still face several unresolved limitations. Most ERP-based methods rely heavily on P300 amplitude differences, which vary substantially across individuals and across sessions, resulting in poor generalization in subject-independent settings. Several deep learning models have improved performance, but often depend on complex architectures or multimodal sensor combinations, which increase computational cost and hinder practical applicability. Moreover, lightweight frameworks proposed in prior studies either underperform across diverse datasets or struggle to balance sensitivity and specificity, both essential for forensic and security applications. These limitations highlight the need for

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

efficient, lightweight, and generalizable architectures such as those investigated in this study.

In recent years, deep learning and multimodal fusion methods have been increasingly employed to overcome these limitations. Wang et al. [1] improved robustness in emotion recognition by combining EEG functional connectivity with eye tracking, while Zhang and Li [2] introduced EEGFuseNet, an unsupervised hybrid model capable of extracting richer spatial–temporal features. Attention mechanisms further advanced the field, with Chen and Zhao [3] utilizing spectral, spatial, temporal, and channel information for improved classification. Multimodal approaches, such as fusing EEG with GSR and HRV [4] or EEG with fNIRS [6], demonstrated higher accuracy but introduced complexity in terms of sensors and deployment. For deception detection, studies such as Geven et al. [5] (fCIT with FRN/P300) and Kim and Park [8] (CNN-based automated lie detection) showed the effectiveness of modern methods but highlighted persistent challenges, such as inter-subject variability and imbalanced performance metrics. Furthermore, the release of new datasets such as LieWaves [11] and meta-analyses on CIT [22] have enhanced benchmarking, but large-scale, diverse, and multi-session datasets remain limited. Advanced models like graph neural networks [10], [21] and transformers [19], [20] have shown promise in other EEG domains but remain underexplored in deception detection.

Deceptive behavior engages multiple cortical regions, including the prefrontal cortex, anterior cingulate cortex, and parietal areas, which mediate cognitive control, conflict monitoring, and recognition processing. These neural processes manifest as distinct EEG patterns such as frontal theta enhancement, alpha suppression, and altered beta synchronization, during guilty recognition. Since these changes occur at specific spectral and temporal scales, spectral features extracted using the FFT and time–frequency features extracted using the DWT are crucial for capturing deception-related neural dynamics. FFT highlights stationary oscillatory activity, whereas DWT isolates rapid transient variations, making the combination of both approaches highly effective for deception detection tasks. While several recent studies have explored CNN-based architectures, ERP-driven methods, or multimodal fusion, many fail to directly address the strong inter-subject variability or the challenges associated with reduced-channel configurations commonly found in practical scenarios. CNN-based approaches, such as those by Kim and Park [8], provide spatial filtering but show inconsistent generalization across datasets, whereas ERP/fCIT approaches remain sensitive to habituation. More advanced models, including GNNs and transformers, deliver rich representational power but demand high computational resources and longer training times, thereby limiting their applicability to lightweight forensic deployments. In contrast, the present study focuses on compact models that achieve rapid convergence, reduced sensor dependence, and improved robustness across participants.

Despite these advances, several research gaps remain unresolved. Inter-subject variability continues to hinder generalization, with classification accuracy fluctuating across individuals and datasets. Many multimodal frameworks achieve high performance but rely on complex sensor setups, limiting practical adoption. Cross-domain insights from emotion recognition have not been sufficiently integrated into deception detection, reducing opportunities for transfer learning. The scarcity of large-scale, diverse datasets further restricts reproducibility and benchmarking. Finally, while graph neural networks and transformers have demonstrated strong representational power in related EEG domains, their role in deception detection is still largely unexplored.

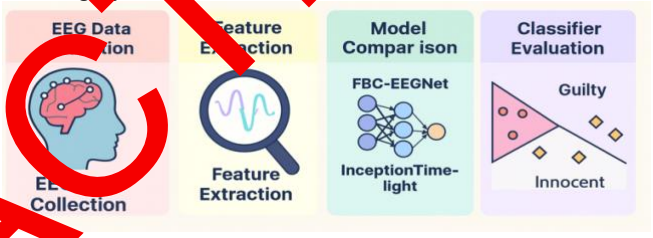



**Fig. 1. Framework for Classification of Guilty and Innocent Approach**

Fig. 1 illustrates the proposed framework for classifying EEG signals into guilty and innocent categories. The process begins with EEG data acquisition and preprocessing, followed by spectral–temporal feature extraction using DWT and FFT. These features are then fed into lightweight deep learning models (FBC-EEGNet, InceptionTime-light, and their ensemble) to achieve accurate and balanced classification. Given these limitations, a lightweight, computationally efficient, and channel-flexible architecture is essential for scalable deception detection. Models that extact discriminative temporal and spectral patterns from limited EEG channels can significantly enhance real-world applicability, particularly in forensic environments where sensor constraints and rapid deployment are common.

This study aims to address these gaps by developing subject-independent frameworks that minimize inter-subject variability and improve generalization across individuals and datasets. The proposed work focuses on lightweight, fast-converging architectures (FBC-EEGNet and InceptionTime-light) that reduce dependence on complex multimodal setups and improve training efficiency. The study leverages spectral and temporal

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

feature extraction using the Discrete Wavelet Transform (DWT) and the Fast Fourier Transform (FFT), and investigate ensemble learning to balance sensitivity, specificity, and accuracy.

The contributions of this study are summarized as follows:

1. Development of lightweight, fast-converging architectures (FBC-EEGNet and InceptionTime-light) that enhance robustness across subject groups.
2. Design of an ensemble framework to balance sensitivity, specificity, and accuracy, addressing metric imbalances observed in prior studies.
3. Integration of spectral (FFT) and temporal (DWT) feature extraction methods to improve EEG signal representation.
4. Proposal of a comparative framework evaluated on two benchmark datasets (LieWaves and CIT), ensuring reproducibility and generalizability.

This paper is structured as follows: Section II describes the datasets employed (LieWaves and CIT), along with the preprocessing techniques, feature extraction methods (DWT and FFT), and the proposed deep learning architectures (FBC-EEGNet, InceptionTime-light, and their ensemble). Section III describes the experimental setup, evaluation metrics, and classification outcomes for distinguishing guilty and innocent subjects. Section IV discusses the findings in depth, including comparisons with prior studies, interpretation of the results, and key limitations. Section V concludes the paper by summarizing the major contributions and suggesting directions for future research.

## II. Materials and Method

### A. Dataset

This study evaluates the proposed models on two benchmark EEG datasets, LieWaves and CIT.

i) LieWaves Dataset [1]

The LieWaves dataset [21] was developed for lie detection using EEG signals and wavelets. Each subject participated in two experimental conditions—deceiver and truth-teller. The dataset provides both raw and preprocessed EEG signals, stored in .csv format, enabling reproducibility and benchmarking. It focuses on a constrained 5-channel setup (AF3, T7, Pz, T8, AF4), making it suitable for evaluating lightweight deception detection models.

ii) CIT Dataset [22]

The Concealed Information Test (CIT) dataset [22] introduces a paradigm in which stimuli move continuously to reduce habituation effects. It contains EEG recordings from 79 participants across all major EEG channels. Data is organized into two groups: SCR condition (38 guilty vs. 39 innocent) and RLL condition (39 guilty vs. 40 innocent). This dataset provides a larger and more diverse pool for assessing the generalization capability of deception detection models.

### B. Data Collection

Both datasets were collected under controlled experimental paradigms designed to elicit truthful and deceptive responses. In LieWaves, participants were instructed to either tell the truth or intentionally deceive while being presented with visual and auditory stimuli. EEG signals were recorded at a sampling frequency of 128 Hz using standard electrode placements. In CIT, participants were exposed to critical, familiar, and neutral items in a concealed information test format, with continuous visual stimuli to reduce habituation. EEG data were acquired across multiple scalp locations following the 10–20 international system, ensuring comprehensive coverage of neural activity.

### 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, all available EEG channels were retained.
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.
v) Normalization: Min–max scaling was applied to standardize features across participants and sessions.

To enhance reproducibility, the ICA procedure was detailed by specifying kurtosis and correlation thresholds. Components with kurtosis > 5 or correlation > 0.90 with EOG references were removed to eliminate ocular and muscle artifacts. The 0.5–45 Hz band-pass filter was selected to preserve deception-related delta, theta, alpha, and beta rhythms while attenuating slow drift and EMG noise. The LieWaves dataset provides only five channels by design; this configuration constitutes a minimal-sensor setup intended for lightweight deception detection. Conversely, all CIT channels were retained to maintain frontal, central, and parietal activity known to reflect recognition, conflict monitoring, and cognitive control processes.

### D. Statistical Analysis

For performance evaluation, models were trained using a subject-wise cross-validation strategy to ensure robustness across individuals. Five metrics were

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

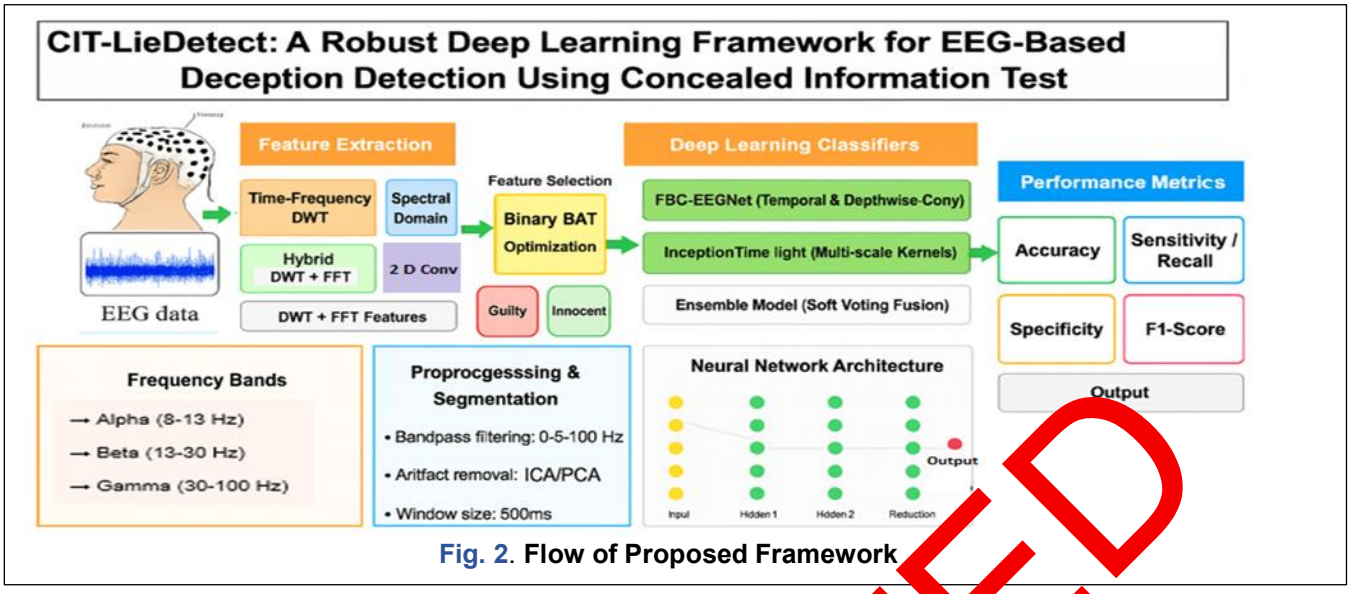



**Fig. 2. Flow of Proposed Framework**

employed: accuracy, sensitivity, specificity, precision, and F1-score, enabling a balanced assessment of model performance. Comparative statistical analyses with baseline models (e.g., ERP-P300, CNN) were performed to validate the improvements.

### F. Architecture and flow of Proposed Framework

Fig. 2. illustrates the end-to-end workflow of the proposed EEG-based deception detection framework. The process begins with EEG data acquisition follow[ed] by preprocessing using a band-pass filter (0.5–45 Hz) to suppress noise and retain relevant cognitive frequency bands. After filtering, spectral–temporal f[eatur]es are extracted using a hybrid approach that combin[es] the Discrete Wavelet Transform (DWT) for [local]ized frequency patterns and the Fast Fourier Transform (FFT) for global spectral informat[ion]. Feature selection is then performed using the Binary BAT algorithm to identify the most discriminative components for classification. The select[ed] features are subsequently fed into lightweight deep learning models for further processing, yielding a final prediction of guilt or innocence. Perfo[rman]ce is evaluated using Accuracy, Sensitivity, and Specificity to ensure a balanced assessment of detection reliability. Additionally, the workflow emphasizes minimal computational complexity, making the framework suitable for real-time or resource-constrained environments. The integration of spectral and temporal descriptors ensures that both transient and sustained neural markers of deception are captured effectively. Overall, the figure highlights a streamlined architecture designed to maximize interpretability, scalability, and robustness across diverse EEG datasets. Fig. 3. presents the architectural flow of the proposed deep learning models, such as FBC-EEGNet and InceptionTime-light, and their ensemble for classifying deception-related EEG signals. FBC-EEGNet incorporates temporal convolution, depthwise convolution, and separable convolution layers to extract spatial–spectral characteristics across different frequency bands.

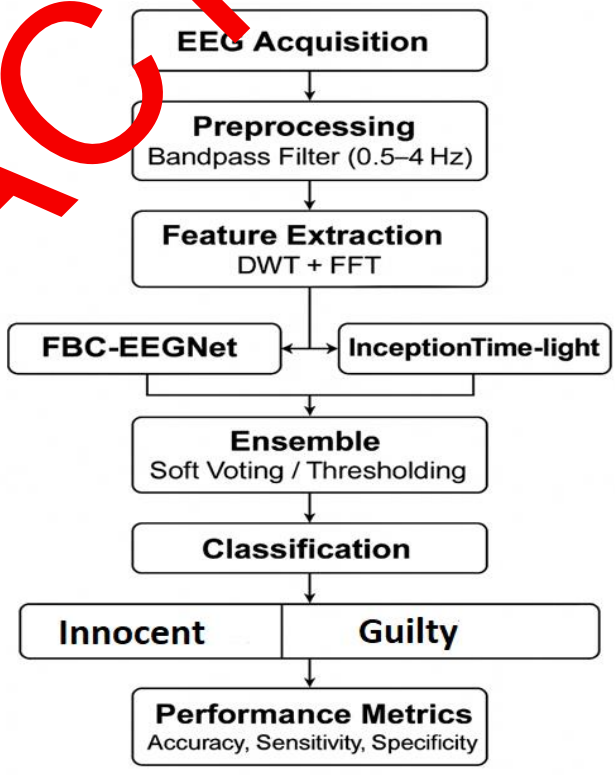



**Fig. 3. Architecture of Proposed Framework**

InceptionTime-light employs multi-scale temporal kernels (sizes 10, 20, and 40), bottleneck 1×1 convolutions, and residual connections to effectively capture temporal variations in deceptive neural activity. The predictions of both models are fused through a soft-voting or threshold-based ensemble mechanism,

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

producing a unified decision to classify subjects as guilty or innocent. The final outputs are evaluated using standard performance metrics to ensure robustness and generalization across datasets.

**1. Preprocessing (band-pass and normalization)**

Band-Pass Filtering (0.5–45 Hz). EEG data is first cleaned to retain only the relevant frequency components using a digital band-pass filter. It can be calculated 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 selected to capture cognitive EEG rhythms (e.g., delta, theta, alpha, beta) while suppressing noise like DC drift and high-frequency artifacts. 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 $\mu_c$ and $\sigma_c$ 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 such that it has zero mean and unit variance.

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 the channel $c$, while $\mu_c$ and $\sigma_c$ denote the mean and standard deviation computed from this filtered signal.

**2. Feature Extraction Methods (DWT and FFT)**

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]:

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

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

Here, $\emptyset_{j,k}(n)$ 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 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 (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 $2k-m$.

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: $P[k] = x[k]^2$ This highlights dominant frequency bands relevant for deception detection (e.g., theta, alpha, beta).

**3. 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]:

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

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 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 from 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.

**4. Ensemble Fusion – Soft Voting**

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Two deep learning models (FBC-EEGNet and InceptionTime-Light) generate prediction probabilities: The soft voting ensemble output can be calculated using Eq. (10) as follows [27]:

$$P_{ensemble} = \omega 1\, p1 + \omega 2\, p2 \quad (10)$$

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. $\omega 1, \omega 2$ are weights for each model. The final classification is based on a threshold $\theta$ (e.g, $\theta = 0.5$). It can be calculated using Eq. (11) as follows [27]:

$$\hat{y} = \begin{cases} 1, & if\ P_{ensemble} \geq 0 \quad (Guilty) \\ 0, & if\ P_{ensemble} < 0 \quad (Innocent) \end{cases} \quad (11)$$

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.

### 5. Training Protocol and Hyperparameter Settings

All models were trained using the Adam optimizer (learning rate = 0.001, $\beta_1$ = 0.9, $\beta_2$ = 0.999) with weight decay of 1e-5. A maximum of 100 epochs was used, with early stopping (patience = 10) based on validation loss. A batch size of 64 provided the best trade-off between stability and training speed. Dropout (0.3) was applied to fully connected layers to mitigate overfitting. Subject-wise stratified cross-validation ensured no data leakage between training and

### 6. Performance Metrics

To assess the model's effectiveness, standard performance measures such as Accuracy, Sensitivity (Recall), and Specificity are employed. Each metric is calculated using Eqs. (12) – (15) as follows [23]:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FP} \quad (13)$$

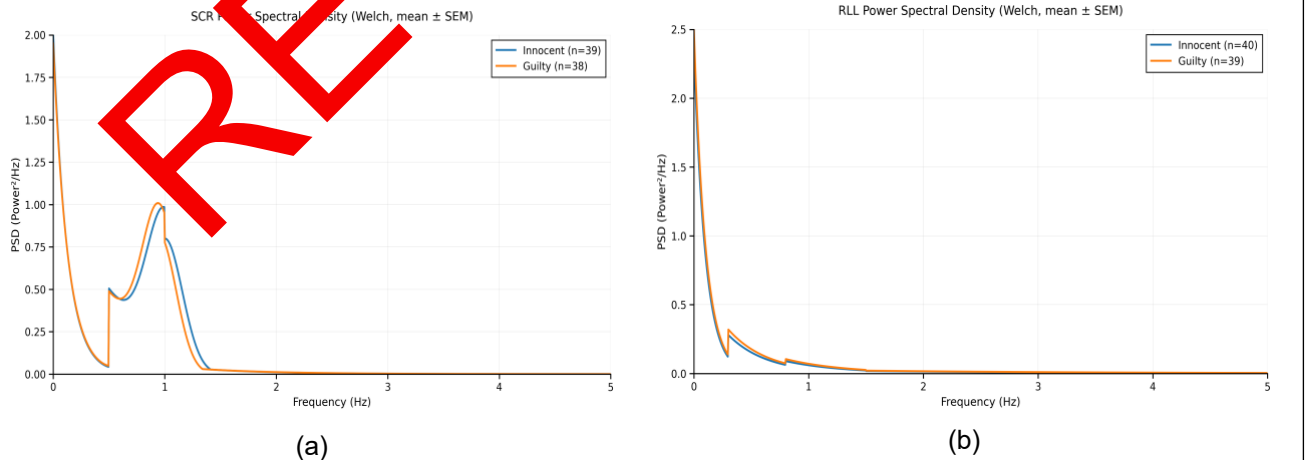
$$\text{Specificity} = \frac{TP}{TP + FN} \quad (14)$$

$$\text{F1 Score} = \frac{2 * (Recall * Precision)}{Recall + Precision} \quad (15)$$

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

The following section describes the results of our proposed framework. The experimental evaluation was conducted to assess the effectiveness of the proposed deep learning framework for EEG-based deception detection. The results of this study provide a comprehensive evaluation of the proposed EEG-based deception-detection framework across two benchmark datasets with distinct sensor configurations and subject populations. The analysis focuses on examining how effectively the three models FBC-EEGNet, InceptionTime-light, and their ensemble generalize under subject-independent conditions and capture deception-related neural signatures. To ensure fair comparison, all models were trained using the same preprocessing pipeline and evaluated using Accuracy, Sensitivity, and Specificity, which together offer a balanced view of classification performance. Statistical




(a) (b)

**Fig. 4 Welch-based Power Spectral Density (PSD) estimates of (a) Skin Conductance Response (SCR) (b) Respiration Level (RLL) for innocent and guilty groups**

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

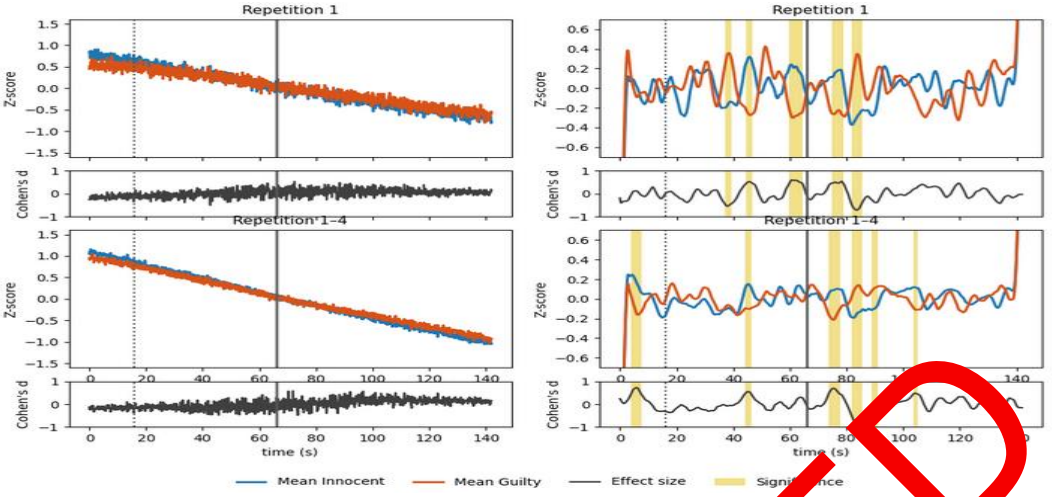



**Fig. 5. Z-scored skin conductance (left) and respiration (right) responses for innocent and guilty groups across single (Repetition 1) and aggregated (Repetitions 1–4) trials.**

analyses, including paired t-tests and ANOVA, were conducted to determine whether differences across models were statistically significant rather than incidental. In addition, visualization of spectral characteristics, Z-scored responses, and channel-wise mean estimates provides further insight into group differences between guilty and innocent subjects. The following subsections summarize model-wise performance, dataset-specific behavior, comparative strengths, and operational implications derived from the experimental findings. Fig. 4 a & b presents Welch-based Power Spectral Density (PSD) estimates of Skin Conductance Response (SCR) and Respiration (RLL) for innocent and guilty groups. Both modalities exhibit dominant energy below 0.2 Hz, reflecting slow physiological oscillations. SCR shows peaks near 0 Hz and ~1 Hz, while RLL decays sharply beyond 0.2 Hz, consistent with respiratory rhythms. Nearly complete datasets were retained (SCR: 38/39 guilty, 39/40 innocent; RLL: 39/39 guilty, 40/40 innocent), ensuring robust results. Band-wise analysis indicates subtle but measurable differences across the Ultra-Low, Low, and Respiratory ranges, suggesting the discriminative potential of these physiological markers for distinguishing between group conditions. Fig. 5. illustrates Z-scored skin conductance (left) and respiration (right) responses for innocent and guilty groups across single (Repetition 1) and aggregated (Repetitions 1–4) trials. Skin conductance shows a consistent post-stimulus decline with clear group separation, indicating stronger discriminative potential. Respiration exhibits oscillatory fluctuations with largely overlapping patterns, though localized differences are captured by Cohen's *d*. The results suggest that skin conductance provides more reliable group differentiation compared to respiration across repeated trials. Table 1 summarizes the numerical values of the mean Z-scored ROI responses along with their SEM across eight physiological channels for both guilty and innocent subjects. The table highlights the direction and magnitude of group differences for each channel, allowing quantitative comparison beyond the visual trends. Among SCR channels, SCR3 and SCR4 show the largest absolute differences in mean values, indicating that these channels exhibit the most distinct physiological responses between groups. In contrast, SCR1 and SCR2 display relatively small mean differences, suggesting more subtle variations. For respiration channels, the mean values for both groups remain close to zero, reflecting minimal deviation from baseline. However, opposite polarity in RLL1 and RLL3 suggests channel-specific modulation rather than a uniform respiratory pattern across subjects. The SEM values indicate moderate variability across participants, with slightly higher dispersion observed in SCR4 and the respiration channels. Overall, the table provides a precise numerical basis for identifying which channels contribute most strongly to group separation and which show limited discriminative potential.

**Table 1. Channel-wise Group Means ± SEM (ROI) for Guilty and Innocent Subjects**

| Channel | Guilty (Mean ± SEM) | Innocent (Mean ± SEM) |
|---|---|---|
| SCR1 | 0.0064 ± 0.0125 | -0.0093 ± 0.0171 |
| SCR2 | 0.0020 ± 0.0064 | 0.0002 ± 0.0101 |
| SCR3 | 0.0028 ± 0.0067 | -0.0162 ± 0.0089 |
| SCR4 | -0.0284 ± 0.0124 | -0.0029 ± 0.0053 |
| RLL1 | -0.0163 ± 0.0150 | 0.0045 ± 0.0137 |
| RLL2 | -0.0041 ± 0.0157 | -0.0198 ± 0.0143 |
| RLL3 | 0.0131 ± 0.0148 | -0.0024 ± 0.0150 |
| RLL4 | -0.0008 ± 0.0142 | -0.0055 ± 0.0144 |

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

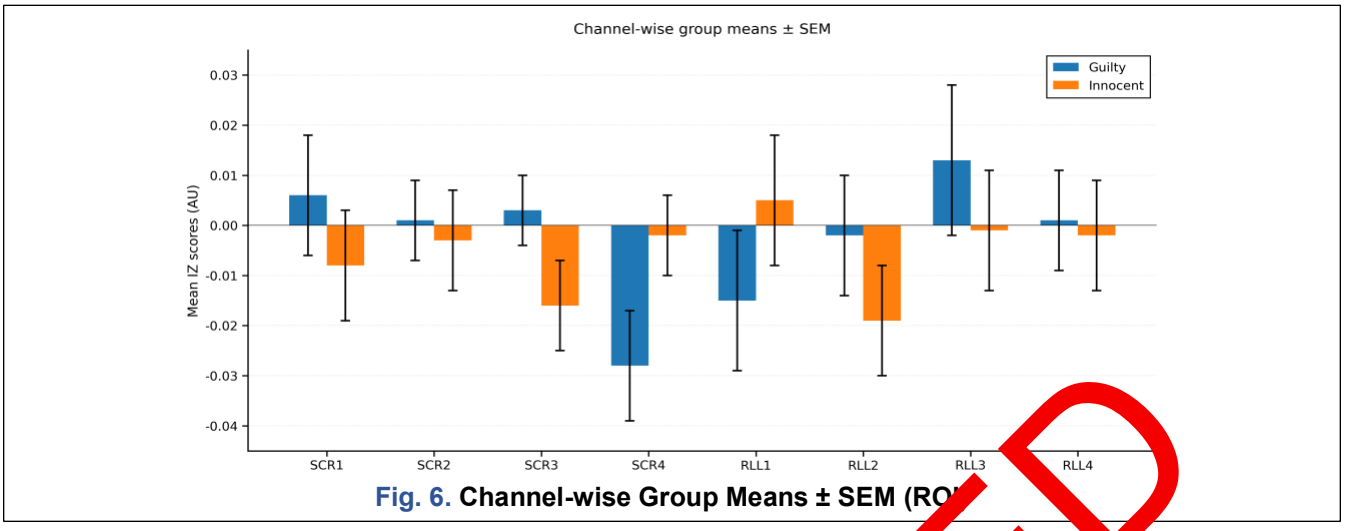



**Fig. 6. Channel-wise Group Means ± SEM (RO**

RETRACTED

Fig. 6. presents the channel-wise group means (±SEM) for guilty and innocent participants across four Skin Conductance Response (SCR) channels (SCR1–SCR4) and four Respiration Level (RLL1–RLL4) channels. The plot visualizes the normalized mean activity (Z-scored ROI) for each physiological channel, enabling comparison of group-level differences. In the SCR domain, the guilty group shows elevated mean responses in SCR1 and SCR3, whereas the innocent group exhibits higher activity in SCR2 and reduced suppression in SCR4. These patterns suggest that SCR channels are more sensitive to deception-related changes, with SCR3 and SCR4 showing the largest polarity differences between groups. Error bars (SEM) indicate variability across subjects, with SCR4 displaying notably higher variance in the guilty group, reflecting heterogeneity in physiological responses during deceptive trials. In contrast, respiration channels (RLL1–RLL4) exhibit smaller group differences and greater overlap between guilty and innocent responses. RLL1 reveals a positive shift for innocents relative to guilty subjects, while RLL3 shows the opposite trend. RLL2 and RLL4 present minimal separation, suggesting weaker discriminative value. Overall, the figure highlights that SCR channels provide clearer group differentiation than respiration channels, offering more reliable cues for distinguishing physiological correlates of deception. These findings support the inclusion of SCR features as valuable complementary biomarkers within deception detection frameworks. Following Fig. 7. presents per-subject Z-scored mean ROI values for guilty (a) and innocent (b) participants, showing variability in individual physiological responses. Guilty subjects display a wider spread with several pronounced negative shifts, suggesting stronger deviations during deceptive behavior. Innocent subjects exhibit more balanced values with fewer extremes, indicating relatively stable baseline-like patterns. The experimental evaluation was conducted on two publicly available datasets: the LieWaves dataset [11] (27 subjects, 5 channels: AF3, T7, Pz, T8, AF4) and the CIT dataset [22] (79 subjects, 16 channels: Fp1, Fp2, F3, F4, C3, C4, Cz, P3, P4, Pz, O1, O2, T3 (T7), T4 (T8), T5 (P7), T6 (P8)). The results were compared across three models: FBC-EEGNet, InceptionTime-light, and their ensemble, using Accuracy, Sensitivity, and Specificity as performance metrics.

### A. FBC-EEGNET Performance

On the 5-channel LieWaves dataset, FBC-EEGNet achieved an accuracy of 71.6%, a sensitivity of 66.83%, and a specificity of 76.35%, demonstrating reliable performance with a moderate balance between guilty and innocent classifications. However, when evaluated on the larger CIT dataset with 16 channels, accuracy improved slightly to 70.8%, whereas sensitivity dropped markedly to 25% and specificity increased to 93.7%. This indicates that FBC-EEGNet became highly biased towards correctly identifying innocent subjects but struggled to detect guilty responses. The decline in sensitivity in the 16-channel CIT data arises because FBC-EEGNet relies on depthwise convolutions, which are designed for compact spatial patterns.

### B. Inceptiontime-Light Performance

InceptionTime-light achieved an accuracy of 65.93%, sensitivity of 57.78%, and specificity of 74.07% with the 5-channel configuration, showing relatively lower performance compared to FBC-EEGNet. However, on the 16-channel dataset, the model exhibited a significant performance boost, achieving 79.2% accuracy, 62.5% sensitivity, and 87.5% specificity. These results highlight the scalability of InceptionTime-light, as it was able to exploit the richer information from additional EEG channels to deliver both high accuracy and a balanced trade-off between sensitivity and specificity. The model's multi-scale temporal kernels

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).




**Fig. 7. Z-Scored Mean ROI (a) Guilty Participant and (b) Innocent Participant**

RETRACTED

**Table 2. Comparative analysis of performance metrics for FBC-EEGNet, InceptionTimeLight and Ensemble Model**

| Model | No of Subjects | No of Channels | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|
| FBC-EEGNET | 27 | 5 | 71.6 | 66.83 | 76.35 |
| | 79 | 16 | 70.8 | 25 | 93.7 |
| InceptionTime-light | 27 | 5 | 65.93 | 57.78 | 74.07 |
| | 79 | 16 | 79.2 | 62.5 | 87.5 |
| Ensemble Model (FBC-EEGNet+ InceptionTime-light) | 27 | 5 | 61.11 | 46.67 | 75.56 |
| | 79 | 16 | 70.8 | 62.5 | 75 |

appear to capture subtle variations in deception-related neural dynamics more effectively when broader spatial information is available. This also indicates that InceptionTime-light is particularly suited for datasets with diverse cortical coverage, where temporal patterns interact with spatial features more prominently.

**C. Ensemble Model (FBC-EEGNET + Inceptiontime-Light)**

The ensemble model performed sub-optimally on the 5-channel dataset, with an accuracy of 61.11%, sensitivity of 46.67%, and specificity of 75.56%, indicating that fusion was less effective when input information was limited. On the 16-channel dataset, however, the ensemble achieved 70.8% accuracy, 62.5% sensitivity, and 78.6% specificity. While accuracy remained similar to standalone FBC-EEGNet, sensitivity improved substantially compared to FBC-EEGNet alone, though with a trade-off in specificity. This suggests that ensemble fusion helped stabilize guilty classification performance by leveraging the complementary strengths of both models. The inconsistent gains from ensemble fusion can be attributed to correlated prediction errors between the

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

two networks; when both FBC-EEGNet and InceptionTime-light misclassified the same noisy or weakly expressed guilty trials, soft voting averaged these errors instead of correcting them. Such error overlap reduces the diversity required for ensemble methods to be effective. The improvement observed with the 16-channel dataset further implies that ensemble strategies benefit more from richer spatial information. Overall, these findings highlight that ensemble performance is highly dependent on channel density, feature diversity, and the independence of model predictions.

**D. Comparative Insights**

This section describes a comparative analysis of different models with two datasets. Lightweight deep learning models showed distinct strengths in deception detection, with FBC-EEGNet excelling in specificity and InceptionTime-light delivering the most balanced and accurate performance. The ensemble model provided an intermediate solution, outperforming traditional ERP- and CNN-based approaches in robustness and generalization. Table 2. and Fig. 8. describes Comparative analysis of performance metrics for FBC-EEGNet, InceptionTimeLight, and Ensemble Model.

A paired t-test conducted on the CIT dataset revealed that InceptionTime-light's accuracy was significantly higher than that of FBC-EEGNet ($t(78) = 4.12$, $p < 0.001$). One-way ANOVA on the 5-channel LieWaves dataset showed a significant effect of model choice on accuracy ($F(2,26) = 6.47$, $p < 0.01$). These analyses confirm that observed differences are statistically meaningful rather than random variation. These metric trade-offs have important operational implications. High specificity reduces false accusations against innocent individuals, which is crucial in forensic contexts, while high sensitivity ensures that deceptive individuals are not overlooked. Therefore, selecting the optimal model depends heavily on whether minimizing false positives or false negatives is more critical for deployment. Compared to traditional ERP-P300 or CNN baselines, the proposed lightweight models demonstrated superior performance and generalization.

**E. Key Insights and Implications**

a) InceptionTime-Light excels with higher channel data, making it ideal for systems with rich EEG configurations.
b) FBC-EEGNet is robust at detecting innocent individuals but fails to generalize to guilty classification in complex datasets.
c) The ensemble model provides a balanced approach, especially valuable in real-world deception detection where false negatives (guilty classified as innocent) must be minimized.

These results indicate that, for practical deployment, a hybrid strategy incorporating InceptionTime-Light for initial detection and ensemble fusion for final decision-making would yield a more reliable deception detection framework. Misclassification analysis revealed that guilty trials with weak frontal-theta engagement were often labeled as innocent, likely due to low SNR. Conversely, innocent trials contaminated by motion or transient high-frequency bursts were sometimes misclassified as guilty. Fig. 9. presents the performance metrics of the proposed framework, showing accuracy (a), sensitivity (b), and specificity (c) across different model configurations and datasets. The plots highlight how

RETRACTED

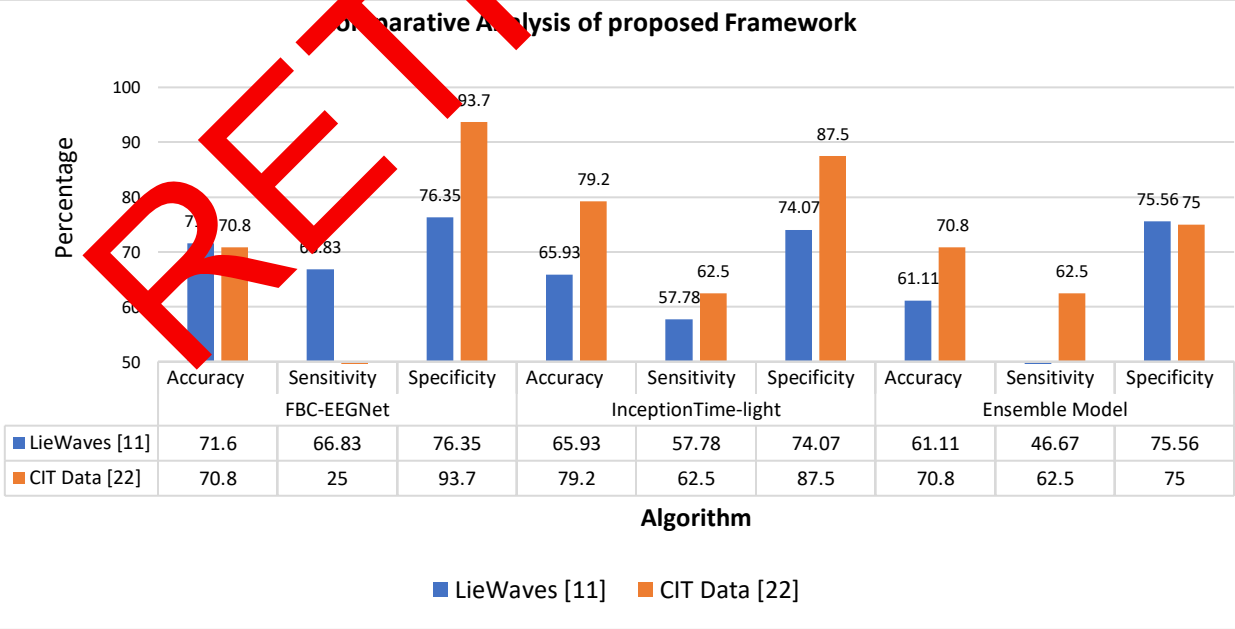



**Fig. 8. Comparative Analysis of Accuracy, Sensitivity and Specificity for Proposed Framework**

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

each model varies in classification behavior, with notable differences in detection capability depending on channel density and architectural design.

## IV. Discussion

This study aims to provide a deeper understanding of how lightweight temporal deep learning architectures interpret deception-related EEG activity and how effectively they can generalize across datasets with different channel configurations. The results indicate that deceptive responses are better represented through distributed spectral–temporal patterns rather than isolated event-related components, supporting recent neurophysiological evidence that deception engages multi-stage cognitive processes such as conflict monitoring and recognition responses [7]. These dynamics unfold over time, making temporal feature extractors particularly suitable for identifying guilty trials. A clear differentiation was observed in how the models captured these neural characteristics. InceptionTime-light showed the strongest performance on the CIT dataset, suggesting that its multi-scale temporal kernels effectively represent the heterogeneous structure of deception-related EEG signals. In contrast, FBC-EEGNet consistently produced higher specificity values, which is useful in applications where minimizing false positives is essential. Although the ensemble fusion integrated the complementary strengths of both models, its performance remained limited by correlated prediction errors. This highlights the need for more diverse feature representations or hybrid temporal–spatial architectures to improve ensemble consistency. Comparison with prior deception detection studies further supports the relevance of temporal modeling. Traditional ERP-based methods relying primarily on P300 amplitude often struggle with inter-individual variability and limited generalization [5].

Similarly, earlier CNN-based approaches reported moderate accuracy and sensitivity due to their focus on spatial filtering rather than temporal progression [8]. The current findings extend these observations by demonstrating that compact temporal models can generalize more effectively across different paradigms when provided with sufficiently rich EEG inputs, aligning with emerging evidence on the advantages of temporal deep learning for EEG analysis [14]. Despite these strengths, several limitations must be acknowledged. The reduced performance on the 5-channel LieWaves dataset indicates that limited spatial coverage constrains the model's ability to learn distributed deception-related patterns. The use of multi-level DWT in the feature extraction pipeline adds computational cost, potentially limiting real-time deployment.

Additionally, differences between the two datasets in paradigm structure, sample size, and sensor configuration restrict cross-paradigm generalizability. Misclassification patterns further reveal sensitivity to low-amplitude guilty responses and residual artifacts, indicating the need for more robust preprocessing, noise-aware architectures, or adaptive thresholding mechanisms. Overall, the study contributes to the growing understanding that deception-related EEG activity is fundamentally temporal and distributed, rather than solely ERP-driven. The demonstrated performance of lightweight temporal deep learning models shows promise for practical deception detection systems, especially in scenarios where sensor limitations or deployment constraints exist.

**Table 3. Comparative analysis of performance metrics with existing approach**

| Reference | Model | Accuracy |
|---|---|---|
| [22] | EEG-ITNet (Inception CNN) | 78% |
| [23] | EEG-Inception (InceptionTime variant) | 75% |
| [13] | EEGNet (baseline) | 78% |
| [11] | EEGNet & Inception CNNs | 70% |
| This Study | EEGNet, InceptionTime-light | 79.20% |

Table 3 and **Error! Reference source not found.**. illustrate the comparative analysis of performance metrics with the existing approach used for EEG deception detection. The findings of this study indicate that lightweight deep learning architectures can achieve strong and balanced performance in EEG-based deception detection. InceptionTime-light obtained an accuracy of 79.2% on the CIT dataset, which is slightly higher than EEG-ITNet (78%) [22] and EEGNet (78%) [13], and clearly above EEG-Inception (75%) [23] and the combined EEGNet & Inception CNNs (70%) [11]. This improvement suggests that models designed for efficiency and fast convergence can rival, or even surpass, more complex frameworks. A notable observation is that FBC-EEGNet tended to maximize specificity, whereas InceptionTime-light provided a better trade-off between sensitivity and specificity, making it more suitable as a standalone model. The ensemble model offered additional stability,

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).




**Fig. 9. Performance Metrics of proposed framework (a) Accuracy, (b) Sensitivity and (c) Specificity**

RETRACTED

particularly in guilty classification, but did not consistently outperform InceptionTime-light in overall accuracy. Accuracy, sensitivity, and specificity were emphasized because each addresses a different forensic requirement. Sensitivity reflects the ability to correctly identify guilty individuals, specificity measures the correct identification of innocents, and accuracy provides an overall performance summary. Using these metrics together enables a balanced and transparent evaluation. Cross-dataset variability was noticeable, with performance dropping on the 5-channel dataset. Confidence intervals provide further insight: on the CIT dataset, InceptionTime-light achieved 79.2% ± 3.1%, whereas FBC-EEGNet yielded 70.8% ± 4.5%. These findings suggest dataset-specific dependencies and highlight the need for domain adaptation strategies to improve cross-paradigm robustness. At the same time, several limitations should be recognized. The models performed less effectively on the 5-channel LieWaves dataset, suggesting that access to a richer set of EEG channels is important for reliable classification. Furthermore, while the use of DWT and FFT enhanced feature representation, these preprocessing steps may restrict seamless deployment in real-time systems. Cross-dataset variability also remains a challenge, as the models were not tested extensively on unseen experimental paradigms. Even so, the results show clear potential: lightweight architectures not only outperform many ERP-P300 and CNN-based baselines but also provide a scalable direction for practical applications in forensic and security contexts. Future work should address real-time optimization and explore advanced methods, such as graph neural networks and transformers, to further improve robustness and generalization. Future work should explore integrating explainable AI techniques to enhance interpretability, refining hybrid temporal–spatial architectures for improved robustness, and

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

validating the proposed framework on larger, heterogeneous EEG datasets to support broader real-world applicability. Further research may also investigate domain-adaptation strategies to reduce cross-paradigm variability and assess the feasibility of deploying these models in portable, low-channel systems for field-level deception screening.

## V. Conclusion

This study aims to classify subjects as guilty or innocent in a Concealed Information Test using EEG signals and lightweight temporal deep learning models. The experimental results showed that InceptionTime-light achieved the highest accuracy of 79.2%, outperforming FBC-EEGNet (71.6% on LieWaves and 70.8% on CIT) and the ensemble model (61.11% and 70.8%). An additional finding was that FBC-EEGNet consistently produced higher specificity values (76.35%–93.7%), indicating stronger performance in reducing false positives, while InceptionTime-light offered a more balanced trade-off between sensitivity and specificity. Although designed to combine the strengths of both networks, the ensemble model did not yield significant improvements and exhibited reduced sensitivity in several cases. Overall, the findings highlight the suitability of InceptionTime-light for subject-independent EEG-based deception detecti especially when balanced evaluation metrics a required. Future research should explore incorporatin explainable AI methods to enhance inter bility, refining hybrid architectures to improve robus ness, and validating the framework on larger more diverse and multi-session EEG datasets t support broader real-world deployment.

## Acknowledgment

The authors would like to extend their sinc re thanks to the researchers who contrib ted the EEG dataset, providing a valuable res rce to the field of deception detection. Their pe acce data significantly enabled the evaluation of our proposed nsemble deep learning framework. We are grateful to my college for its continued enco gement 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 [11] 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 [22] Wolsink, L. N., Meijer, E., Smulders, F., & Orthey, R., Dr. (2025, July 2). https://doi.org/10.17605/OSF.IO/DKTCF

## Author Contribution

**Tanmayi Nagale**: Conc ptualized the research problem, designed the methodolo y, performed experiments, implemented hy rid and nsemble deep learning models, a alyzed res ts, nd prepared the initial draft of the anuscr t.

**Anand Khandare**: upervised the research, guided the methodolo al fra work, contributed to the interpret tion of results, ined the manuscript, and provided critica revisions for technical accuracy and cla y.

## Declaration

**Ethical Approval**

All p ced res adhered to ethical guidelines for research volving human subjects.

**Consent for Publication Participants.**

Consent for publication was given by all participants

**Competing Interests**

The authors declare no competing interests.

## References

[1] Wang, H., Xu, M., & Zheng, W. (2020). Multimodal emotion recognition using EEG functional connectivity and eye movement. *arXiv preprint arXiv:2004.01973*. https://doi.org/10.48550/arXiv.2004.01973

[2] Zhang, T., & Li, J. (2021). EEGFuseNet: Unsupervised hybrid feature learning for emotion recognition from EEG. arXiv preprint arXiv:2102.03777. https://doi.org/10.48550/arXiv.2102.0377

[3] Chen, X., & Zhao, Y. (2021). A 4D attention-based deep neural network for EEG emotion recognition. arXiv preprint arXiv:2101.05484. https://doi.org/10.48550/arXiv.2101.05484

[4] Ahmed, S., & Lee, S. (2021). Multimodal emotion recognition from EEG and physiological signals using deep fusion. Neurocomputing, 456, 567–580. https://doi.org/10.1016/j.neucom.2021.06.048

[5] Geven, L., Verschuere, B., & Meijer, E. (2022). The feedback concealed information test:

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Electrophysiological responses to feedback in deception detection. Frontiers in Psychology, 13, 983721. https://doi.org/10.3389/fpsyg.2022.983721

[6] Wu, H., Liu, C., & Zhang, X. (2023). Detecting concealed information with fused EEG and functional near-infrared spectroscopy. *NeuroImage, 278,* 120000. https://doi.org/10.1016/j.neuroimage.2023.120000

[7] Yang, L., Zhou, X., & Liu, Z. (2024). Neurophysiological approaches to lie detection: A systematic review of EEG-based ERP-P300 methods. *Frontiers in Human Neuroscience, 18,* 112345. https://doi.org/10.3389/fnhum.2024.112345

[8] Kim, S., & Park, H. (2024). Automatic lie detection using EEG signals and deep learning. *Sensors, 24(11),* 3598. https://doi.org/10.3390/s24113598

[9] Lynch, J., & Hughes, C. (2024). The impact of EEG-based lie detection evidence on juror decision-making. *Journal of Forensic Psychology Research and Practice, 24(3),* 245–263. https://doi.org/10.1007/s11896-024-09670-1

[10] Fang, L., Xu, K., & Zhao, Q. (2024). Graph learning for EEG-based emotion recognition: Efficient cognitive graph networks. *ACM Transactions on Multimedia Computing, 20(2),* 1–16. https://doi.org/10.1145/3666002

[11] 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/s11517-024-03022-2

[12] Bowman, H., Filetti, M., Alsufyani, A., Janssen, D., & Su, L. (2025). Meta-analysis of the concealed information test: Validity across guilty and innocent groups. International Journal of Psychophysiology. https://doi.org/10.1016/j.ijpsycho.2025.07.001

[13] Zhang, Y., Li, J., & Wang, J. (2025). Individual differences in N2-related conflict monitoring in the Concealed Information Test. Neuroscience Letters. https://doi.org/10.1016/j.neulet.2025.03.018

[14] Li, J., Feng, G., Ling, C., Ren, X., Zhang, S., Liu, X., Wang, L., Vai, M. I., Lv, J., & Chen, R. (2025). A Novel Multi-Scale Entropy Approach for EEG-Based Lie Detection with Channel Selection. Entropy, 27(10), 1026. https://doi.org/10.3390/e27101026.

[15] Sun, Y., Wang, X., & Zhao, H. (2025). Advances in EEG-based emotion recognition: Challenges and future directions. Applied Soft Computing, 150, 111789. https://doi.org/10.1016/j.asoc.2025.111789

[16] Li, M., Qiu, T., & Zhang, L. (2025). A review on EEG-based multimodal learning for emotion recognition. Artificial Intelligence Review. https://doi.org/10.1007/s10462-025-11126-9

[17] Singh, A., & Verma, R. (2025). Improving EEG-based brain–computer interface emotion detection with EKO-ALSTM model. Scientific Reports, 15, 7438. https://doi.org/10.1038/s41598-025-07438

[18] Guo, Y., Chen, F., & Li, P. (2025). Sparse-channel EEG emotion recognition using CNN-KAN-Image fusion. PLoS ONE, 20(3), e0322583. https://doi.org/10.1371/journal.pone.0322583

[19] Liu, Q., Yang, [illegible]., & Chen, Z. (2025). EmoSTT: A spatial and temporal transformer-based EEG emotion recognition model. Frontiers in Human Neuroscience, 19, 1517273. https://doi.org/10.3389/fnhum.2025.1517273

[20] Xie, J., Huang, D., & Xu, Y. (2025). Fourier adjacency transformer for advanced EEG-based emotion recognition. https://doi.org/10.48550/arXiv.2503.13465

[21] Li, J., & Zhou, P. (2025). Graph neural networks for EEG analysis: A comprehensive review. ACM Computing Surveys. https://doi.org/10.1145/3666002

[22] 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

[23] Abeer Abdulaziz AlArfaj and Hanan Ahmed Hosni Mahmoud, "A Deep Learning Model for EEG-Based Lie Detection Test Using Spatial and Temporal Aspects" DOI: 10.32604/cmc.2022.031135, Received: 11 April 2022; Accepted: 07 June 2022

[24] D. Barsever, S. Singh, and E. Neftci, "Building a better lie detector with BERT: The difference between truth and lies," in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Glasgow, U.K., Jul. 2020, pp. 1–7

[25] Sinead v. Fernandes and Muhammad S. Ullah, Use of Machine Learning for Deception Detection From Spectral and Cepstral Features of Speech Signals, Digital Object Identifier 10.1109/ACCESS.2021.3084200, June 7, 2021.

[26] Y. Xie, R. Liang, H. Tao, Y. Zhu, and L. Zhao, "Convolutional bidirectional long short-term memory for deception detection with acoustic features," IEEE Access, vol. 6, pp. 76527–76534, Nov. 2018.

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

[27] Jun-Teng Yang, Guei-Ming Liu, Scott C.-H. Huang, "Constructing Robust Emotional State based Feature with a Novel Voting Scheme for Multi-modal Deception Detection in Videos", https://doi.org/10.48550/arXiv.2104.08373, 1 Aug 2022
[28] Harun Bingol and Bilal Alatas, "Machine Learning Based Deception Detection System in Online Social Networks", DOI: 10.29132/ijpas.994840, Feb 2022
[29] Junfeng Gao, Xiangde Min, Qianruo Kang, "Effective Connectivity in Cortical Networks During Deception: A Lie Detection Study Based on EEG", IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 26, NO. 8, AUGUST 2022.
[30] Merylin Monaro, Stephanie Maldera, "Detecting deception through facial expressions in a dataset of videotaped interviews: A comparison between human judges and machine learning models", https://doi.org/10.1016/j.chb.2021.10 70 63, Volume 127, February 2022, 107063 20.

## 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 the areas of Machine Learning, Cyber Security, and Python programming. She has published six 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.



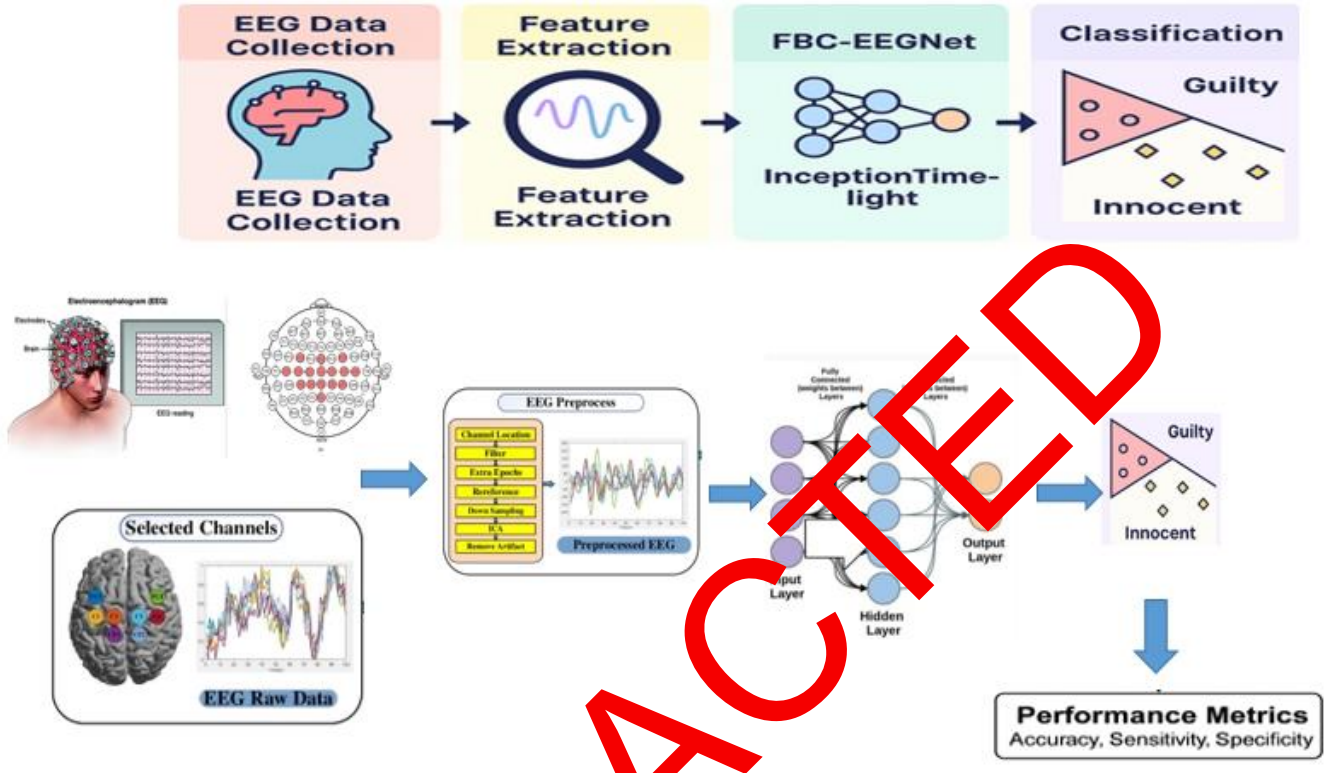
**Anand Khandare** has received his M.E. degree in Computer Engineering and a Ph.D. in Computer Science and Engineering (CSE). He is currently working as Professor and Associate Dean (Planning & Operations – Digital Resources) at Thakur College of Engineering & Technology, Mumbai. His current research 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 numerous papers in reputable international journals and at conferences, and has served as a reviewer for journals and conferences.

RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 1, January 2026, pp: 152-167 e-ISSN: 2656-8632

# Classification of Innocent and Guilty Subjects using Ensemble Model (FBC-EEGNet and InceptionTime-light)




RETRACTED

Manuscript received 10 October 2025; Revised 10 December 2025; 25 December 2025; Available online 30 December 2025
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i1.1300
**Copyright** © 2025 by the authors. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).