Article | Received 21 February 2026; Revised 9 April 2026; Accepted 8 May 2026; Published 12 June 2026
https://doi.org/10.55092/neuroelectronics20260007

Band-specific EEG connectivity and graph-theoretical features for attention detection

Sami Al Majanini[1], Mohamad Khalil[2,*] and Wassim El Falou[2]

[1] Department of Electronic Engineering, Lebanese University (UL), Tripoli, Lebanon
[2] Faculty of Engineering and Doctoral School for Sciences and Technology, Lebanese University (UL), Tripoli, Lebanon

* Correspondence author; E-mail: mohamad.khalil@ul.edu.lb.

Highlights:

- Subject-dependent MI classification achieves up to Accuracy = 97.23%, F1 = 0.970, AUC = 0.998 with Logistic Regression.
- MI-based broadband EEG features + Logistic Regression reach Accuracy = 87.91% ± 6.34%; F1 = 0.893, AUC = 0.927 LOSO accuracy.
- Theta-band small-world-like topology during Attention contrasts with hub-dominated, segregated alpha-band networks during Inattention.
- Beta-band networks show the richest pattern: higher global efficiency, clustering, and node strength during Attention across Pearson, COH, and MI measures.

Abstract: Electroencephalogram based attention detection becomes growing research filed in past years because his enormous application from education to clinical diagnosis. Exiting approaches to detect attention has often relied on features extracted from individual electrodes, such as frequency or entropy measures. In this study we examined graph-theoretical features derived from Coherence, Pearson correlation, and Mutual Information across theta, alpha, beta, and broadband bands in ten participants. Nine graph features capturing network integration and hub centrality were extracted per epoch, and false discovery rate correction identified 6 significant features in both theta, alpha bands and 9 in beta band, with notable effect sizes (Cohen's d = 0.88–1.70). Theta networks showed small-world-like topology during Attention, alpha networks were more segregated during Inattention, and beta networks exhibited the richest differences. Subject-dependent classification reached 97.23% accuracy (Mutual information, Logistic Regression, beta band; F1-score = 0.970, Area under curve = 0.998), while leave one subject out confirmed broadband MI features as the most consistent (87.91% ± 6.34%; F1 = 0.893, AUC = 0.927), though higher inter-subject variability highlights instability was noted, which is expected in this small group of subjects. These results suggest that combining linear and information-theoretic connectivity measures captures complementary aspects of attentional networks, with broadband MI features outperforming traditional electrode-level approaches and offering a promising approach for EEG-based cognitive monitoring.

Keywords: EEG-based attention detection; graph-theoretical analysis; functional connectivity; mutual information connectivity; coherence and Pearson correlation

1. Introduction

Attention is an essential cognitive process that allows us to focus on relevant stimuli while filtering out distractions, enabling efficient interaction with our environment. In educational settings, attention is a must for effective learning and knowledge retention. The ability to measure and classify attention levels has numerous applications in various domains, including e-learning [1], driving [2], and clinical diagnostics [3]. Beyond traditional behavioral and self-report measures like watching eye gaze or head pose. Electroencephalography (EEG), offer a non-invasive technique for understanding attention states by capturing neural responses to sensory inputs [4]. Prior EEG studies have demonstrated that attention-related processes are strongly reflected in specific frequency bands particularly theta (4–8 Hz), alpha (8–13 Hz), and beta (13–30 Hz) which are associated with cognitive control, sensory inhibition, and task engagement, respectively [5].

Based on these findings, many studies have focused on extracting features from these frequency bands using various signal-processing approaches, including time-domain features, frequency-domain features, and entropy-based measures [6]. For example, Enrique *et al.* [7] employed time-domain features such as zero crossings, skewness, and the Hurst exponent to detect driver distraction. Vortmann *et al.* [8] utilized power spectral density (PSD) features extracted from alpha, beta, theta, and gamma bands to classify attention levels in an augmented reality paradigm. Similarly, Fan *et al.* [9] applied entropy-based features, including Shannon entropy, approximate entropy, and sample entropy, to identify driver distraction and fatigue states.

Despite these advances, attention detection faces several problems like extracting electrode wise features, such as spectral or entropy-based measures, which neglecting the truth that attention is a network-level interactions [10]. When functional connectivity is considered, several studies use a single measure without examining whether linear and nonlinear metrics capture complementary aspects of brain dynamics. In addition, to the best of our knowledge there is no single unified research that investigate the contribution of each EEG band in the classification accuracy. Another important limitation concerns data availability: compared to other EEG research domains, publicly available attention datasets remain scarce, with many studies relying on self-collected data, which limits reproducibility and cross-study validation [11]. Finally, most existing approaches depend on high-density EEG systems ($\geq$ 32 channels), which is impractical for real-world scenarios.

In response to these challenges, we propose a systematic band-specific graph-theoretical framework for EEG-based attention detection. Moving beyond current single electrode features extraction methods, in this framework we model the brain as a network using graph derived from three complementary connectivity measures: Coherence, Pearson correlation, and Mutual Information, capturing linear oscillatory, linear amplitude, and nonlinear statistical dependencies, respectively. Furthermore, instead of aggregating features across frequencies, each band is analyzed independently to characterize how attention state changes across theta, alpha, beta, and broadband ranges. In addition, a low-density eight-channel EEG system was used to enhance practical deployability in applied, everyday scenarios such as classrooms and driving contexts. Given the limited availability of public datasets in this domain, we collected our own dataset, our work also contributes a novel dataset to the field.

Our study offers the following contributions to the field of EEG-based cognitive decoding:

(a) A systematic comparison of three functional connectivity measures Coherence, Mutual Information, and Pearson correlation within a unified graph-theoretical framework.

(b) Independent band-specific analysis across theta, alpha, beta, and broadband ranges, revealing frequency-dependent network signatures of attentional states.

(c) Presents and validates the approach on a self-collected EEG dataset acquired using a practical low-density setup, contributing empirical data to a field where publicly available datasets remain limited.

The next sections of the manuscript are structured as follows. Section 2 outlines our experimental design, EEG recording setup functional connectivity estimation, graph construction, feature extraction and classification. Section 3 details Statistical analysis results and classification performance, Section 4 discusses the findings, limitations, and potential directions for future research and Finaly Section 5 conclude the paper.

2. Methods

2.1. Participants and experimental paradigm

Ten university students aged between 19 and 23 years participated in this study. Ethical compliance was maintained throughout the study, with all participants granting informed consent in accordance with local institutional requirements. Given the current lack of a universally accepted experimental protocol for detecting attention via EEG, a novel audiovisual paradigm was designed to elicit distinct attentive and inattentive cognitive states.

During the Attention condition, participants watched a 10-minute lecture video to induce sustained attentional engagement. During the Inattention condition, participants watched a silent nature video intended to minimize cognitive load and task engagement. This contrast was designed to capture neural dynamics associated with attentional modulation while remaining suitable for low-density EEG acquisition. Data acquisition occurred within a quiet, purpose-built laboratory setting at the Doctoral School of the Lebanese University to reduce environmental noise and motion-related artifacts. An overview proposed research architecture, including EEG recording, functional connectivity computation, graph construction, feature extraction, and classification, is illustrated in Figure 1.

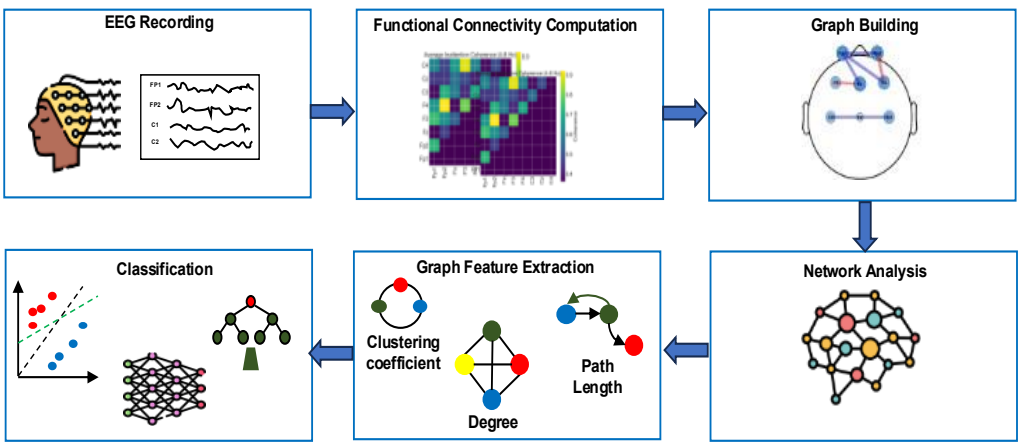

Figure 1. Methodological workflow for EEG-based attention detection.

2.2. EEG acquisition and processing

For EEG acquisition, we utilized an Enobio eight-channel EEG system set to a sampling rate of 500 Hz. We positioned the electrodes according to the international 10–20 system at the following sites: Fp1, Fp2, Fz, F3, F4, C3, Cz, and C4. These channels were selected based on prior studies highlighting the role of frontal and central regions in attention control and executive processing [12].

To prepare our raw EEG signals for analysis, we utilized the MNE-Python library. We first applied a band-pass filter (1–45 Hz) to remove slow drifts and high-frequency noise. Then employed independent component analysis (ICA) to identify and remove components ocular and muscular artifacts.

To perform epoch wise functional connectivity calculation the cleaned EEG signals were segmented into overlapping epochs of 10 seconds duration with 50% overlap, following previous studies [13]. This window length balances temporal resolution with the stability of EEG measures, while the overlap increases the number of training samples for subsequent connectivity and classification analyses. For each participant and condition, this procedure yielded approximately 119 epochs, resulting in data with dimensions (epochs × channels) ≈ (119 × 8).

2.3. Functional connectivity computation

To examine how inter-regional brain communication fluctuates between attention and inattention, we compute functional connectivity (FC) for every 10-second epoch and for each frequency band (Theta: 4–8 Hz, Alpha: 8–13 Hz, Beta: 13–30 Hz, broadband:4–35 Hz) using three complementary measures: Pearson correlation (r), Coherence (COH), and Mutual Information (MI). These measures capture different aspects of neural coupling: linear amplitude relationships (r), frequency-specific oscillatory synchronization (COH), and nonlinear statistical dependencies (MI).

(a) Pearson Correlation

We utilized Pearson correlation to assess the linear relationship between two EEG signals, $X_i(t)$ and $Y_j(t)$, from channels i and j, and is defined as:

$$\mathrm{r_{ij}} = \frac{\mathrm{cov(X_i, X_j)}}{\sigma_{\mathrm{X_i}}\sigma_{\mathrm{X_j}}} \tag{1}$$

Where $cov(X_i, X_j)$ is the covariance and σ denotes the standard deviation. The resulting matrix is symmetric with values ranging from −1 to +1, indicating perfect negative correlation and perfect positive linear correlation respectively. Pearson correlation is computationally simple and widely used in EEG-based attention research, including backward models for auditory attention decoding [14]. Though it has several limitations like capturing only linear associations and sensitivity to volume conduction.

(b) Coherence

We applied Coherence to measures the frequency-domain similarity between two signals, phase and amplitude consistency across time, and is defined as:

$$\mathrm{Coh_{ij}(f)} = \frac{|\,\mathrm{S_{ij}(f)}\,|^2}{\mathrm{S_{ii}(f)} \cdot \mathrm{S_{jj}(f)}} \tag{2}$$

where $S_{ij}(f)$ is the cross-spectral density between channels i and j, and $S_{ii}(f)$, $S_{jj}(f)$ are the power spectral densities. Coherence ranges from 0 to 1, reflecting how tow channels are synchronized in a specific frequency band. It is particularly relevant in attention research, e.g., frontal–parietal theta

coherence has been associated with top-down controlled attention [15]. Limitations include the assumption of signal stationarity within the analysis window.

We compute COH using the multitaper spectral estimation method implemented in MNE-Python via the spectral_connectivity_epochs function. Compared to single-window methods the multitaper approach improves spectral concentration and reduces estimation variance. This procedure was done for each epoch within predefined frequency bands: theta, alpha, beta, and broadband. The multitaper method internally determines the spectral smoothing and frequency resolution based on the epoch duration (10 s) and sampling frequency (500 Hz). All COH values were averaged across frequency bins within each band (faverage = True), resulting in a single connectivity value per channel pair per band per epoch. This process produced approximately 119 coherence matrices per condition for each subject.

(c) Mutual Information

Recognizing that neural interactions are often non-linear, we employed MI to quantifies the shared information between two signals, by capturing both linear and nonlinear dependencies, and is defined as:

$$\mathrm{MI_{ij}} = \sum_{x,y} \mathrm{p_{ij}}\,(x,y)\log\left(\frac{\mathrm{p_{ij}}(x,y)}{\mathrm{p_i}(x)\mathrm{p_j}(y)}\right) \quad (3)$$

where $p_i(x)$ and $p_j(y)$ are the marginal distributions, and $p_{ij}(x,y)$ is the joint distribution. MI is symmetric, non-negative, and higher values indicate stronger dependence. It is particularly suited to identify complex interactions that correlation or coherence may miss, like relationships between EEG and peripheral autonomic signals during high-workload tasks [16].

We compute MI independently for each epoch to capture nonlinear statistical dependencies between EEG channels and we estimate it using the k-nearest neighbors (k-NN) approach implemented in mutual_info_regression from scikit-learn, with the number of neighbors set to $k = 5$, with one channel as input and the other as the target variable. This procedure was repeated for all channel pairs, resulting in symmetric connectivity matrices per epoch. The k-NN estimator avoids assumptions about underlying distributions and is practical for capturing nonlinear interactions in EEG signals.

For each participant and condition, these FC measures generated 8 × 8 connectivity matrices for each epoch and frequency band, which were subsequently used for graph-theoretical feature extraction and classification analyses.

2.4. Graph-theoretical feature extraction

To investigate the topological properties of brain networks during attention and inattention states, we extract a comprehensive graph-theoretical features from each epoch's connectivity matrix using Networkx toolbox in python. For each participant, connectivity type $c \in$ {Pearson r, COH, MI} and frequency band $f \in$ {theta, alpha, beta}, proportional thresholding was applied by retaining the top 40% of strongest edges in each network. This approach will reduce the influence of weak and potentially spurious connections and ensures consistent network density across conditions while preserving meaningful topological structure. Importantly, given the relatively small network size ($N = 8$ nodes), more strict thresholds would result in sparse or disconnected graphs, which will violate the validity of graph-theoretical metrics particularly those dependent on global connectivity such as characteristic path length and global efficiency. Finaly The selected threshold falls within the commonly recommended 30%–50% range and aligns with prior findings highlighting the sensitivity and instability

of graph metrics across threshold levels, as discussed in [17]. Since connectivity weights represent similarity rather than distance, path-length-dependent measures (global efficiency and average path length) computed on an inverted-weight graph where each edge distance was defined as $d_{ij} = 1/w_{ij}$ for $w_{ij} > 0$, ensuring that stronger functional connections correspond to shorter topological distances. Strength-based measures (node strength, clustering coefficient) were computed directly on the original weighted graph.

For each epoch we extract the following graph measures:

- Node Strength: The sum of weights of all edges connected to a node, representing overall connectivity. For a weighted adjacency matrix $W = [w_{ij}]$, the strength of node i is defined as:

$$s_i = \sum_{j=1}^{N} w_{ij} \tag{4}$$

  Network-level statistics were obtained by computing the mean and standard deviation of s_i across all nodes.

- Clustering Coefficient: Quantifies the tendency of a node's neighbours to form tightly interconnected clusters The weighted clustering coefficient of node $i$is defined as:

$$C_i = \frac{1}{k_i(k_i - 1)} \sum_{j,h} (w_{ij} w_{ih} w_{jh})^{1/3} \tag{5}$$

  where $k_i$is the degree of node i. Mean and standard deviation across nodes were used as network descriptors.

- Betweenness Centrality (weighted): measures the proportion of shortest paths passing through a node, reflecting its role as a hub (mean and standard deviation). It is defined as:

$$BC_i = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \tag{6}$$

  where $\sigma_{st}$is the total number of shortest paths between nodes $s$and t, and $\sigma_{st}(i)$is the number of those paths that pass-through node i.

- Path length**:** indicates the number of edges (steps) between two nodes. In a network summary, the Average Path Length (mean and standard deviation) measures the overall efficiency or "closeness" of the network, computed using the distance matrix $D = [d_{ij}]$:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}^{shortest} \tag{7}$$

  where $d_{ij}^{shortest}$is the shortest path length between nodes $i$and j.

- Global Efficiency: Evaluates the capacity of the network for parallel information transfer and is defined as:

$$E_{glob} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}^{shortest}} \tag{8}$$

  This procedure yields a 9-dimensional feature vector per epoch (Node Strength: 2, Clustering Coefficient: 2, Betweenness Centrality: 2, Path length: 2, Global Efficiency: 1).

For each participant, connectivity type, frequency band, and attentional condition, the 119 epochs produced a feature matrix:

$$X_{s,c,f} \in R^{119\times 9} \tag{9}$$

By concatenating the feature matrices of the two attentional conditions (Attention and Inattention) for the same subject, connectivity type, and frequency band, we obtain:

$$X_{s,c,f}^{total} = \begin{bmatrix} X_{s,c,f,Attention} \\ X_{s,c,f,Inattention} \end{bmatrix} \in R^{238\times 9}, Y_{s,c,f} \in K^{238}, K = \{Attention, Inattention\} \tag{10}$$

The selection of these graph measures is motivated by prior studies demonstrating their sensitivity to cognitive fluctuations and network-level attentional processes [17].

2.5. Statistical analysis and classification

To avoid pseudo-replication, we first aggregated feature values across all epochs for each participant, generating a single representative metric per condition. We then employed the non-parametric Wilcoxon signed-rank test to compare these conditions; we prioritized this test over parametric alternatives due to our modest sample size (n = 10) and potential violations of normality in some features. To assess the magnitude of these differences, we calculated Cohen's d effect sizes, which offer a practical perspective on the results independent of the limitations of small-sample statistical power.

Given that we tested nine distinct graph features per connectivity measure, we performed Benjamini-Hochberg False Discovery Rate (FDR) correction [18] to mitigate the risk of Type I errors. We defined each connectivity type (COH, MI, Pearson) as a discrete 'family' of tests for this correction, setting the threshold at $q < 0.05$. Throughout our results, we report both raw and corrected p-values to maintain full transparency, while advising caution when interpreting effect sizes, which can remain volatile in smaller cohorts [19].

For state classification, we evaluated both subject-dependent and subject-independent frameworks. For the subject-dependent approach, we utilized 3-fold cross-validation, training and testing classifiers within each participant's own data. In another hand Subject-independent classification (LOSO) one subject was held out as the test set while the remaining participants were used for training. This procedure was repeated for each subject to assess generalization across individuals.

To prevent data leakage in the LOSO protocol, we applied Principal Component Analysis (PCA) exclusively to the training set within each fold, retaining components explaining 95% of the variance (n_components = 0.95). The same transformation fitted only on training data was then applied to the held-out test subject. This ensures that no information from the test subject influences the feature transformation, preventing a subtle form of leakage that can inflate LOSO accuracy estimates. Given the low feature dimensionality (9 features), PCA was not applied for aggressive dimensionality reduction but solely to orthogonalize the feature space and address inter-feature redundancy identified in the correlation analysis (Section 3.1). The mean number of components retained across folds was approximately 7 out of 9, explaining a cumulative variance of 95%.

Eight machine learning classifiers were trained and evaluated on the feature sets:

- Logistic Regression (LR)
- Stochastic Gradient Descent (SGD)

- Multilayer Perceptron (MLP)
- Random Forest (RF)
- Extra Trees (ET)
- Gradient Boosting (GB)
- XGBoost (XGB)
- Support Vector Machine (SVM)

For each classifier, the hyperparameters utilized during training are detailed in Table 1. To ensure a robust assessment of our framework, we evaluated classification performance across three key metrics: accuracy, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC). These metrics were computed independently for every combination of connectivity type $c \in \{Pearsonr, COH, MI\}$ and frequency band $f \in \{\text{theta}, \text{alpha}, \text{beta}\}$ as well as for the broadband condition. This comprehensive approach ensures unbiased evaluation and provides insight into the contributions of different connectivity measures and frequency bands to attentional state decoding.

Table 1. Hyperparameters used for all eight machine learning classifiers.

Classifier	Key Hyperparameters	Values/Notes
Logistic Regression	penalty, solver, C, max_iter	penalty = 'l2', solver = 'liblinear', C = 1.0, max_iter = 1000
Stochastic Gradient Descent	loss, penalty, alpha, max_iter	loss = 'log', penalty = 'l2', alpha = 0.0001, max iter = 1000
Multilayer Perceptron	Hidden layer sizes, activation, solver, learning rate, max iter	Hidden layer sizes = (100,), activation = 'relu', solver = 'adam', learning rate = 0.001, max iter = 500
Random Forest	Number of estimators, max depth, min samples split, min samples leaf	Number of estimators = 100, max depth = None, min samples split = 2, Min samples leaf = 1
Extra Trees	Number of estimators, max depth, min samples split, min samples leaf	Number of estimators = 100, max depth = None, min samples split = 2, Min samples leaf = 1
Gradient Boosting	Number of estimators, learning rate, max depth, subsample	Number of estimators = 100, learning rate = 0.1, max depth = 3, subsample = 1
XGBoost	Number of estimators, max depth, learning rate, subsample, colsample by tree	Number of estimators = 100, max depth = 3, learning rate = 0.1, subsample = 1.0, colsample by tree = 1.0
SVM	C, kernel, gamma, probability	C = 1.0, kernel = 'rbf', gamma = 'scale', probability = True

3. Results

3.1. Feature correlation analysis

To evaluate the degree of informational redundancy among the extracted graph metrics, we conducted a Spearman rank correlation analysis. We chose the Spearman method specifically to accommodate the non-normal distributions inherent in these topological features (see Section 3.2). For this exploratory analysis, observations were pooled at the epoch level across all subjects to provide a comprehensive assessment of feature interdependencies. We performed these analyses separately for each connectivity estimator (COH, MI, Pearson) and attentional state (Attention, Inattention).

While reporting is focused on the theta band as a representative frequency, we observed that these correlation structures remained qualitatively consistent across the alpha, beta, and broadband conditions. It is important to emphasize that this correlation analysis is purely descriptive; it does not constitute inferential statistical testing. As noted, all formal comparisons between conditions were performed on subject-level aggregated features to ensure independence of observations and prevent pseudo-replication. The resulting lower-triangular Spearman correlation matrices Figure 2 suggest significant multicollinearity across the feature set, with several pairs exhibiting coefficients exceeding $|r| = 0.70$. Notably, in the MI-Attention condition, the mean clustering coefficient and its standard deviation showed a strong positive correlation $r = 0.914$, suggesting they capture overlapping aspects of local segregation. Furthermore, we observed an inverse relationship between node strength and characteristic path length. $r = -0.843$.

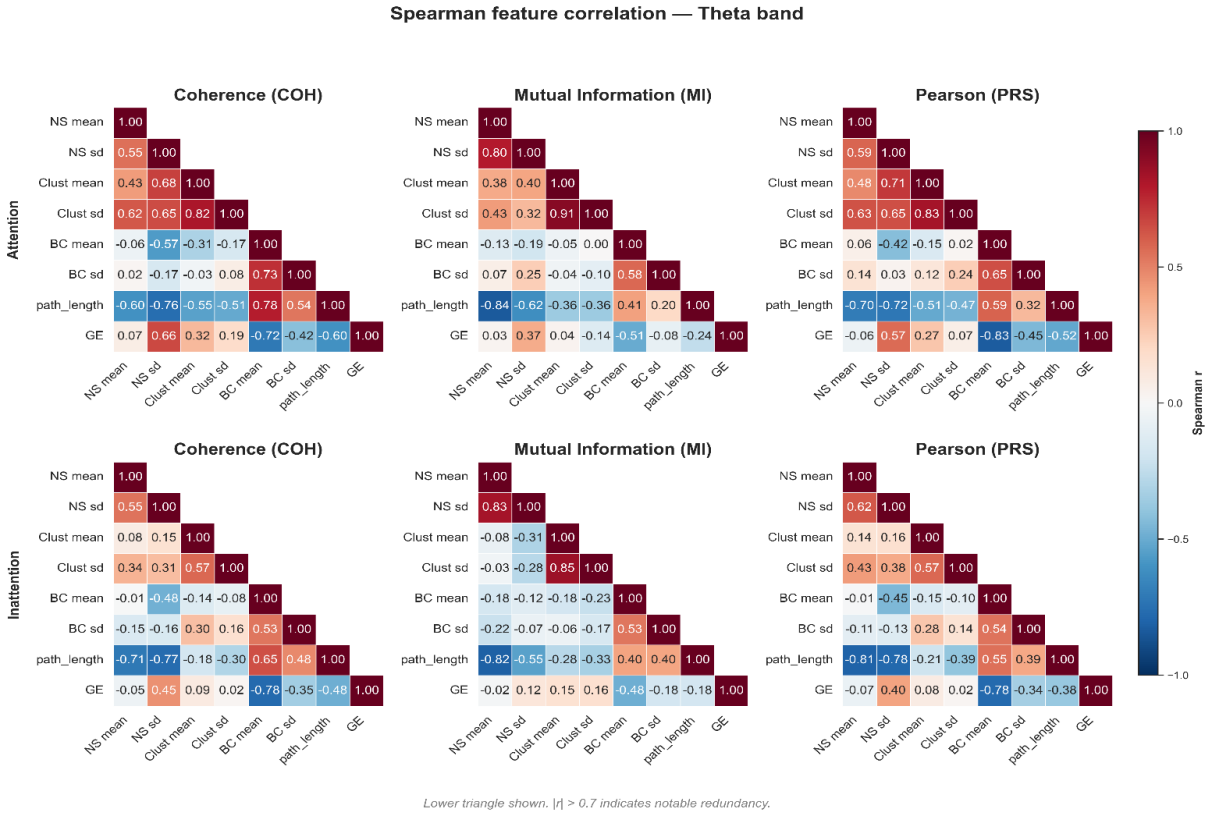

Figure 2. Spearman rank correlation matrices of nine graph-theoretical features.

Despite these high correlations, we choose to retain all nine features for two specific reasons:

(1) Theoretical Heterogeneity: While statistically correlated, the mean and standard deviation of a metric represent distinct neurobiological properties specifically, the global network "average" versus the spatial diversity (nodal heterogeneity) of functional roles across the scalp [20].

(2) Orthogonalization via PCA: we addressed potential for unstable coefficient estimates or overfitting associated with redundant features via PCA. By integrating PCA directly into the machine learning pipeline and applying it within each Leave-One-Subject-Out (LOSO) fold (as detailed in Section 2.5), we transformed the raw, redundant feature space into an orthogonal set of linearly uncorrelated components [21]. This procedure ensures that the pairwise redundancy among raw features is mathematically addressed prior to classification, enhancing the overall robustness and generalizability of our model.

3.2. Graph features statistical results

Although the study followed a paired design, a parametric paired *t*-test was not adopted. This choice was motivated by the small sample size ($N = 10$) and the empirical distributions of the extracted graph features. As illustrated by the violin plots in Figure 3, several features exhibit clear deviations from normality, including skewness and multi-modal patterns, particularly for variability-related measures.

Accordingly, the non-parametric Wilcoxon signed-rank test was used to assess differences between Attention and Inattention conditions. This test is robust to non-normal distributions and suitable for small-sample paired analyses. Statistical significance was set at $p < 0.05$, and effect sizes were quantified using Cohen's d.

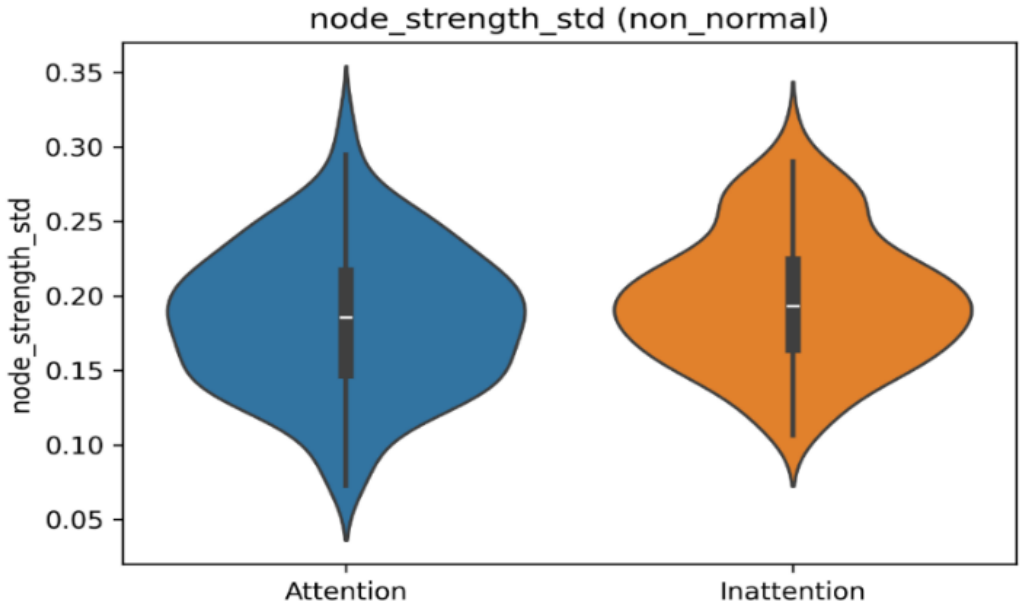

Figure 3. Violin plots showing the distribution of node strength graph feature for Attention and Inattention conditions, motivating the use of non-parametric statistical testing.

3.2.1. Theta-band preliminary network observations during attention and inattention

In the theta band (4–8 Hz), six features survived the FDR correction ($q < 0.05$) within this preliminary ten-participant investigation. Although our sample size limits statistical power, the consistently large effect sizes (Cohen's d = 0.88 to 1.70) suggest these differences are not spurious and warrant investigation in larger samples.

We observed a large-effect-size increase in COH-based node strength mean during Attention (p_raw = 0.0020, p_corrected = 0.009, d = 1.62; Figure 4a), as did Pearson-derived node strength mean (p_raw = 0.0020, p_corrected = 0.018, d = 1.50; Figure 4c).

Furthermore, Pearson-derived clustering coefficient mean significantly increased during Attention (p_raw = 0.0098, p_corrected = 0.029, d = 1.23; Figure 4d), alongside clustering coefficient standard deviation (p_raw = 0.0195, p_corrected = 0.044, d = 0.88; Figure 4e). We hypothesize that indicates enhanced local modularity, likely reflecting the recruitment of specialized cognitive control subnetworks. The higher clustering standard deviation suggests increased nodal heterogeneity, whereby certain channels possibly frontal given their established role in theta-band cognitive control.

may show disproportionately stronger local clustering than others. Both observations carry only moderate effect sizes relative to the other findings.

Both COH- and Pearson-derived path length mean were significantly longer during Inattention (COH: p_raw = 0.0020, p_corrected = 0.009, d = −1.70, Figure 4b; Pearson: p_raw = 0.0059, p_corrected = 0.026, d = −1.26, Figure 4f). These large effect sizes tentatively suggest that the theta-band network may be more segregated during passive viewing, with information requiring more intermediate steps to propagate. The topological coherence between the node strength and path length findings stronger coupling during Attention corresponding to shorter paths is internally consistent and adds plausibility to these preliminary observations.

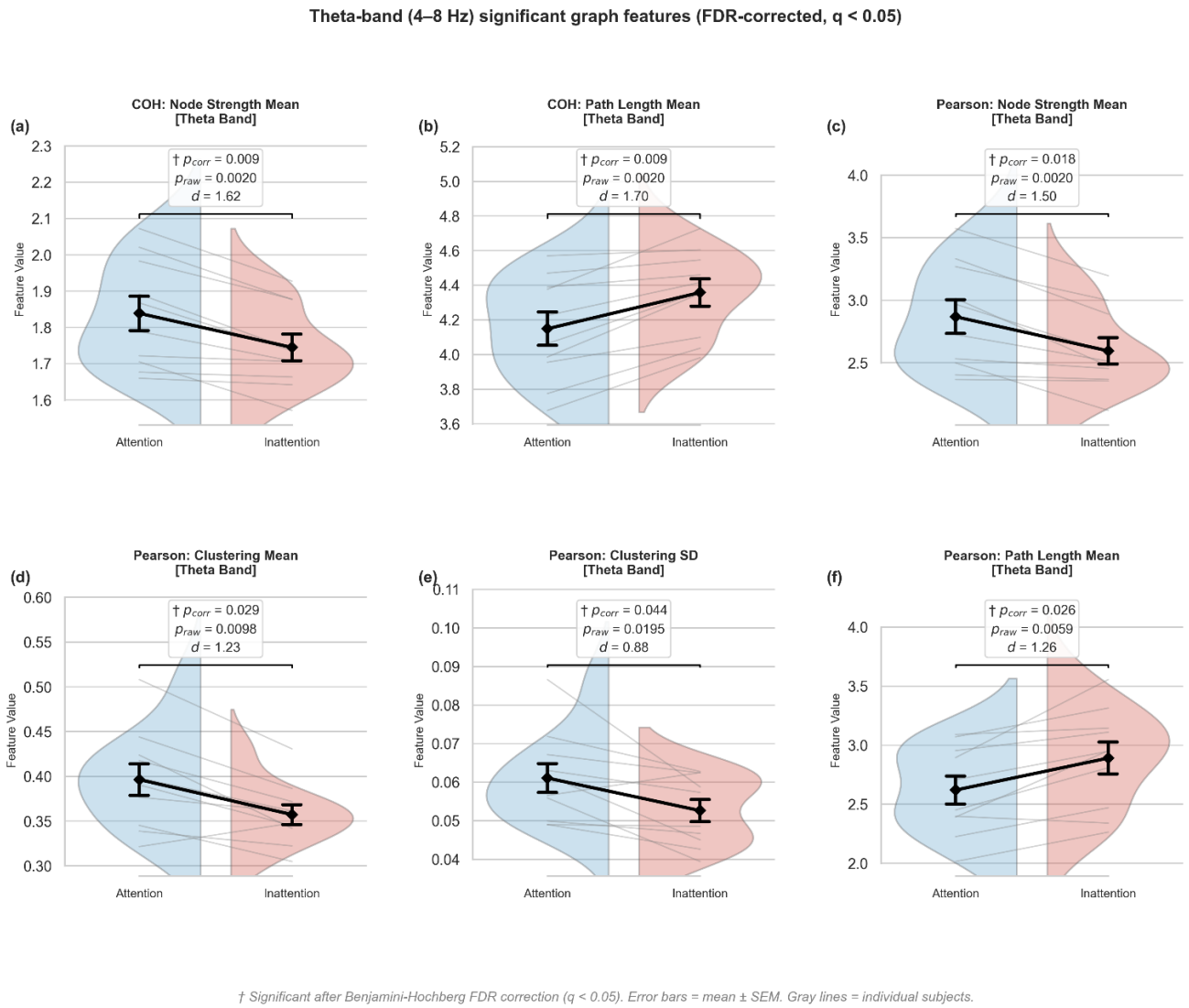

Figure 4. Theta-band (4–8 Hz) graph-theoretical features showing significant differences between Attention and Inattention conditions after Benjamini-Hochberg FDR correction (q < 0.05). Violin plots show the distribution of subject-mean feature values. Error bars indicate mean ± SEM. † FDR-corrected significance. Individual subject trajectories shown as gray lines.

Considered together, the pattern of simultaneously higher clustering and shorter path lengths during Attention is suggestive of a small-world-like topology, which is hypothesized to support optimal cognitive processing through a balance of local specialization and global integration. However, the small sample (N = 10) and the use of a single experimental session per participant, these results should be interpreted as preliminary evidence not an established effect. No MI-based theta features reached significance after FDR correction, suggesting that within this sample and paradigm, linear oscillatory synchronization captured by COH and Pearson correlation appears more sensitive to theta-band attentional modulation than nonlinear MI-based coupling, though this observation also requires replication.

3.2.2. Alpha-band preliminary network observations during attention and inattention

In the alpha band (8–13 Hz), six features survived FDR correction across this ten-participant preliminary sample, with large effect sizes (Cohen's d = 1.05 to 1.62) suggesting within-sample robustness despite the limited statistical power. Unlike the theta band, significant alpha features were distributed across all three connectivity measures including one MI-based feature, tentatively suggesting that both linear and nonlinear dependencies may contribute to alpha-band network differences between conditions, though this observation is based on a single surviving MI feature and should not be over-interpreted.

We observed that COH-based betweenness centrality mean and standard deviation were both higher during Inattention (mean: p_raw = 0.0039, p_corrected = 0.012, d = −1.10, Figure 5a; std: p_raw = 0.0020, p_corrected = 0.012, d = −1.18, Figure 5b). These observations tentatively suggest a tendency toward more hub-dominated, centralized alpha-band network topology during passive viewing.

COH-based path length standard deviation was also significantly higher during Inattention (p_raw = 0.0039, p_corrected = 0.012, d = −1.05, Figure 5c), as was MI-derived path length standard deviation (p_raw = 0.0020, p_corrected = 0.018, d = −1.62, Figure 5d). The higher path length variability during Inattention captured by both linear oscillatory coherence and nonlinear mutual information suggests increased heterogeneity in inter-node communication distances, reflecting a more irregularly structured and segregated network during passive nature-video viewing. Notably, this is the only alpha-band finding where MI reached significance after FDR correction, tentatively suggesting that nonlinear dependencies contribute to capturing path length variability during the inattentive state.

Pearson-derived path length mean (p_raw = 0.0039, p_corrected = 0.018, d = −1.30, Figure 5e) and path length standard deviation (p_raw = 0.0020, p_corrected = 0.018, d = −1.49, Figure 5f) were both higher during Inattention, consistent with the COH and MI path length findings above. Collectively, the path length pattern across three connectivity measures tentatively points toward a more segregated alpha-band network topology during Inattention.

Overall, our alpha-band observations tentatively suggest a contrast between a more distributed, locally integrated network during Attention and a more hub-dominated, segregated network during Inattention.

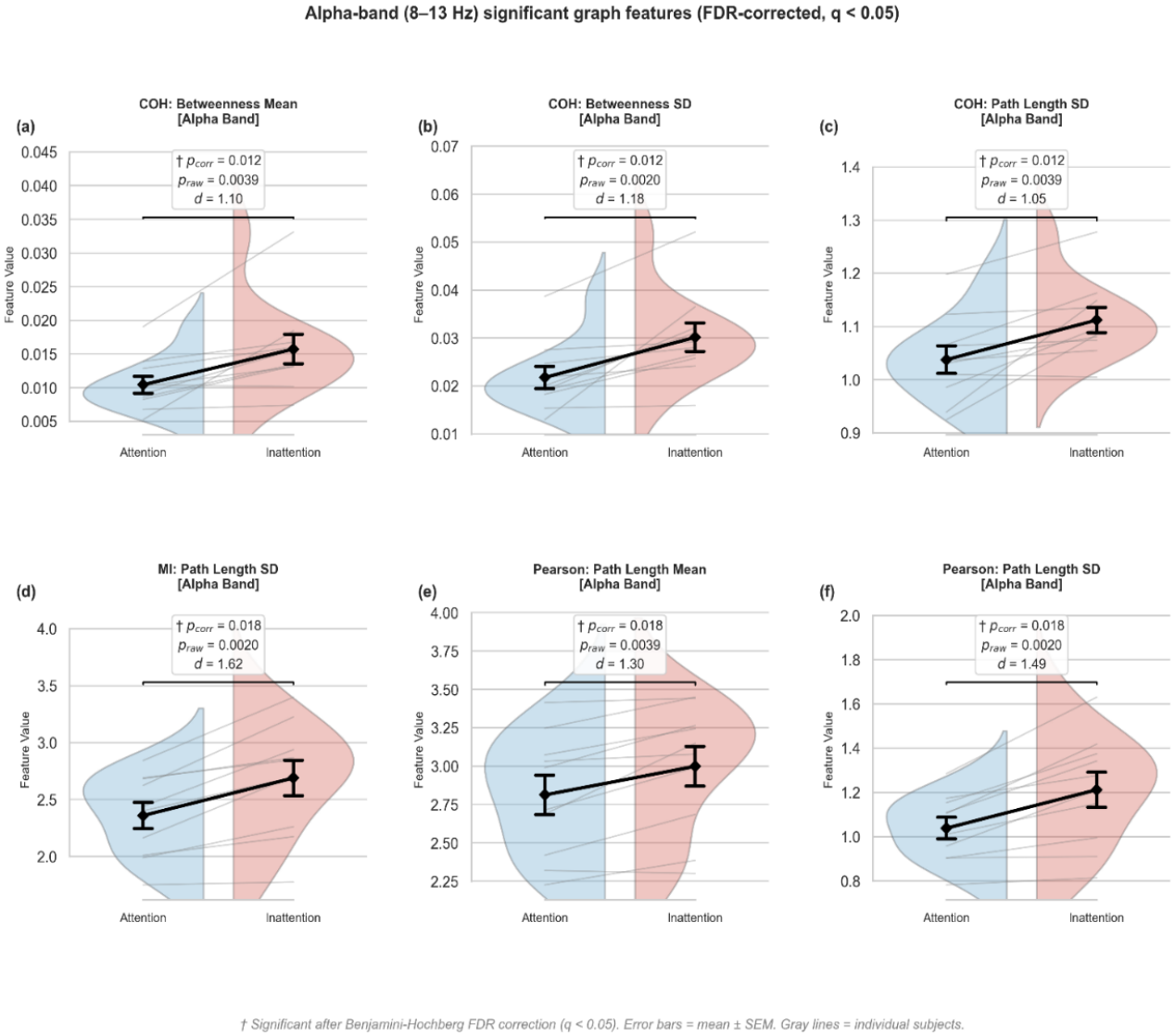

Figure 5. Alpha-band (8–13 Hz) graph-theoretical features showing significant differences after Benjamini-Hochberg FDR correction (q < 0.05). Violin plots show the distribution of subject-mean feature values. Error bars indicate mean ± SEM. † FDR-corrected significance. Individual subject trajectories shown as gray lines.

3.2.3. Beta-band preliminary network observations during sustained attention

In the beta band (13–30 Hz), nine features survived the FDR correction (q < 0.05) across this ten-participant preliminary investigation, representing the richest pattern of significant differences across all three frequency bands examined. Effect sizes ranged from moderate to large (Cohen's d = 0.88 to 1.20), with the largest effects observed for Pearson-derived path length and betweenness features. As with our theta and alpha findings, these results are considered preliminary, given our sample size. We found that path length mean and standard deviation were both significantly longer during Inattention across COH (mean: p_raw = 0.0059, p_corrected = 0.044, d = −1.06, Figure 6a; std: p_raw = 0.0098, p_corrected = 0.044, d = −1.09, Figure 6b) and MI (mean: p_raw = 0.0059, p_corrected = 0.026, d = −1.04, Figure 6c; std: p_raw = 0.0098, p_corrected = 0.029, d = −1.03, Figure 6d) connectivity measures. The convergence of both linear oscillatory and nonlinear measures on longer beta-band path lengths during Inattention

tentatively suggests a marked trend toward more segregated network topology during passive viewing across multiple coupling frameworks.

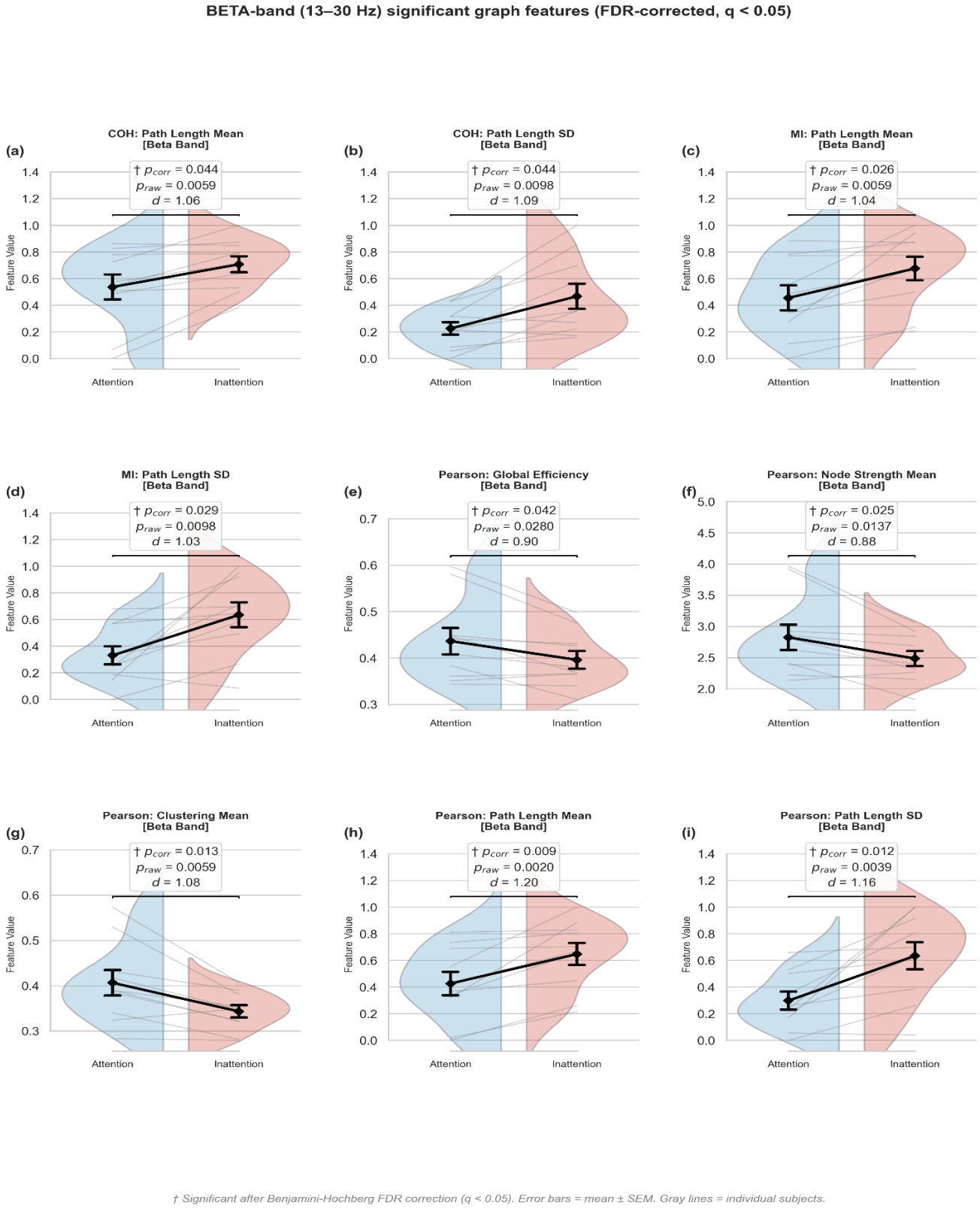

Figure 6. Beta-band (13–30 Hz) graph-theoretical features showing significant differences after FDR correction (q < 0.05). † FDR-corrected significance. Error bars = mean ± SEM. Gray lines = individual subjects.

Conversely, Pearson-derived global efficiency was higher during Attention ($p_{raw} = 0.028$, $p_{corrected} = 0.042$, $d = 0.90$, Figure 6e). Higher global efficiency indicating more efficient parallel information transfer across the beta-band network during sustained lecture viewing.

Pearson-derived node strength mean ($p_{raw} = 0.0137$, $p_{corrected} = 0.025$, $d = 0.88$, Figure 6f) and clustering coefficient mean ($p_{raw} = 0.0059$, $p_{corrected} = 0.013$, $d = 1.08$, Figure 6g) were both higher during Attention. These observations suggest stronger overall coupling and enhanced local modularity during sustained attentional engagement in the beta ban. The combination of higher node strength, higher clustering, and higher global efficiency during Attention in the beta band presents an internally coherent picture of a more integrated, densely coupled network during the lecture condition, though the small sample precludes any strong topological conclusions.

Pearson-derived path length mean ($p_{raw} = 0.0020$, $p_{corrected} = 0.009$, $d = -1.20$, Figure 6h) and path length standard deviation ($p_{raw} = 0.0039$, $p_{corrected} = 0.012$, $d = -1.16$, Figure 6i) were also significantly longer during Inattention, consistent with the COH and MI path length findings above and further supporting the pattern of greater network segregation during passive viewing in the beta band.

Collectively, these beta-band metrics suggest that active attention is characterized by a more integrated and efficient network organization, while inattention is associated with increased segregation.

Across the three frequency bands, our data tentatively suggest dissociable network signatures of attentional states. Theta-band networks showed consistent small-world-like organization during Attention captured by linear measures (COH and Pearson), suggesting oscillatory synchronization plays a primary role in theta-band attentional modulation. Alpha-band networks showed predominantly segregated, hub-dominated topology during Inattention. Beta-band networks showed the richest pattern of significant features. These band-specific signatures align with existing hypotheses that theta, alpha, and beta oscillations serve complementary, frequency-dependent roles in attentional processing [22], though we emphasize that these patterns require future validation in larger cohorts.

3.3. Subject-dependent classification results

We evaluated classification performance across all frequency bands, and EEG connectivity features (COH, Pearson correlation, MI). Each model was assessed using 3-fold cross-validation to ensure reliable performance estimates.

Subject-dependent classification consistently outperformed LOSO, reflecting within-subject optimization. F1-scores and AUC values (Table 2) confirm that accuracy gains are not artefactual. In the beta band, MI features with Logistic Regression achieved the highest accuracy of 97.23% (F1 = 0.970, AUC = 0.998), with alpha-band MI (Random Forest: 92.87%, F1 = 0.929, AUC = 0.975) and Pearson (Logistic Regression: 91.23%, F1 = 0.916, AUC = 0.985) also performing strongly. Beta-band COH features (Random Forest: 88.9%, F1 = 0.895, AUC = 0.898) and theta-band Pearson features (SGD: 89.73%, F1 = 0.882, AUC = 0.905) showed competitive performance. Figure 7, a faceted grouped bar chart, visually compares the performance of each classifier across bands and connectivity types, highlighting that MI features combined with Logistic Regression consistently provide the best within-subject accuracy.

Table 2. Highest average subject-dependent classification accuracies (%) for each EEG frequency band and connectivity measure using multiple classifiers, evaluated with 3-fold cross-validation.

Frequency Band	Connectivity	Best Classifier	Accuracy (%)	F1 Score–AUC
Theta	COH	Extra Trees	85.67%	0.861–0.861
Theta	MI	Logistic Regression	87.42%	0.867–0.884
Theta	**Pearson R**	**SGD**	**89.73%**	**0.882–0.905**
Alpha	COH	SVM	81.2%	0.822–0.830
Alpha	**MI**	**Random Forest**	**92.87%**	**0.929–0.975**
Alpha	Pearson R	Logistic Regression	91.23%	0.916–0.985
Beta	COH	Random Forest	88.9%	0.895–0.898
Beta	**MI**	**Logistic Regression**	**97.23%**	**0.970–0.998**
Beta	Pearson R	Gradient Boosting	92.54%	0.923–0.964

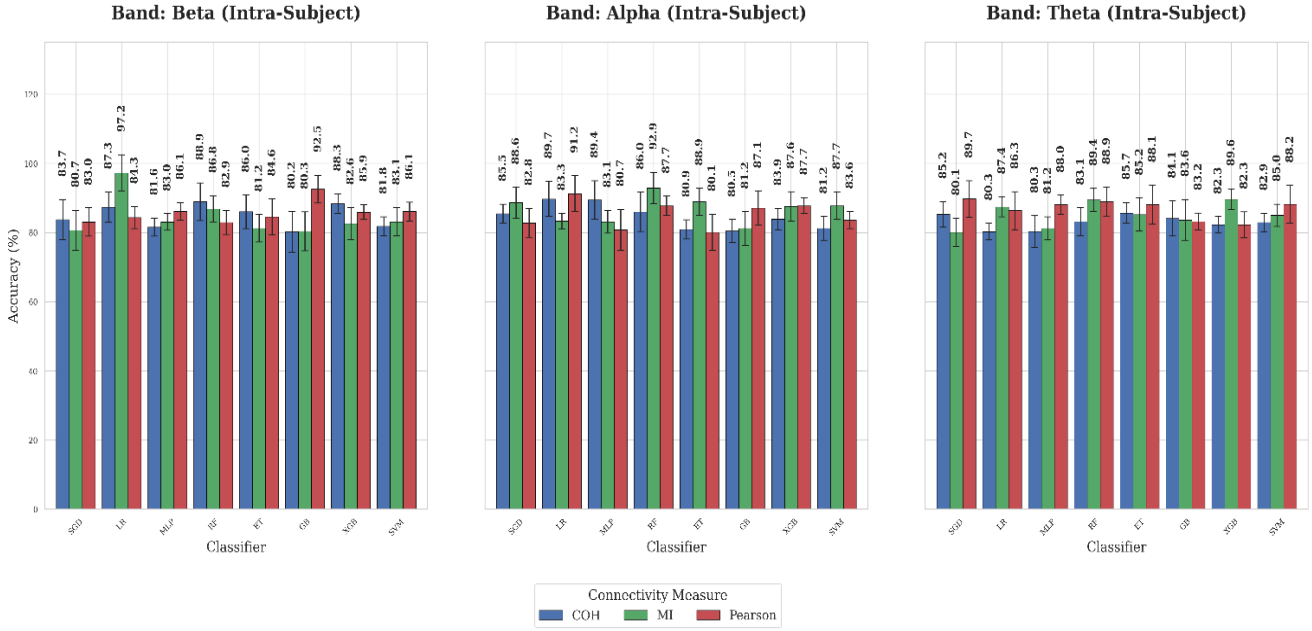

Figure 7. Faceted grouped bar chart showing subject-dependent classification performance across theta, alpha, beta, and broadband EEG features (COH, Pearson correlation, MI) for all classifiers. Bars indicate average accuracy (%) with standard deviation.

3.4. Loso classification result

We also evaluated classification performance using LOSO cross-validation protocol to assess model generalizability. Performance varied depending on the frequency band, connectivity measure, and classifier, as summarized in Table 3. As expected, LOSO performance was lower than the subject-dependent results, underscoring the inherent challenges of cross-subject generalization in EEG-based attention detection.

Despite this Broadband MI features with Logistic Regression achieved the highest LOSO accuracy (87.91% ± 6.34%, F1 = 0.893, AUC = 0.927), with notably lower SD than single-band results, indicating more stable cross-subject generalization. Beta-band COH features (SGD: 76.88% ± 14.69%) and Pearson features (Gradient Boosting: 75.92% ± 19.29%) showed the strongest single-band LOSO performance.

ROC curves for the top-performing LOSO model in each band are presented in Figure 8. The broadband MI Logistic Regression model achieved the highest mean AUC (0.927 ± 0.119), confirming that multi-band integration provides the most robust and generalizable attention detection framework. The large standard deviations in LOSO results indicate substantial inter-subject variability, which is expected given the small sample (N = 10) and the absence of subject-specific calibration in LOSO evaluation. Figure 9, a faceted grouped bar chart, visually compares the performance of each classifier across bands and connectivity types

Table 3. Highest LOSO classification accuracies for each frequency band and connectivity measure across all classifiers. Values represent mean ± standard deviation.

Frequency Band	Connectivity	Best Classifier	Accuracy (%) ± SD	F1 Score–AUC
Theta	COH	SVM	67.10 ± 12.8	0.677–0.716
Theta	MI	SVM	63.36 ± 20.70	0.632–0.652
Theta	**Pearson R**	**SVM**	**71.30 ± 11.92**	**0.712–0.794**
Alpha	COH	Logistic Regression	61.18 ± 11.04	0.611–0.647
Alpha	MI	SGD	63.57 ± 22.9	0.636–0.655
Alpha	**Pearson R**	**MLP**	**69.33 ± 14.76**	**0.696–0.752**
Beta	**COH**	**SGD**	**76.88 ± 14.69**	**0.770–0.785**
Beta	MI	SVM	71.43 ± 12.31	0.726–0.728
Beta	Pearson R	Gradient Boosting	75.92 ± 19.29	0.759–0.796
Broadband	COH	SGD	81.93 ± 12.07	0.821–0.836
Broadband	Pearson R	Extra Tree	84.41 ± 18.22	0.858–0.888
Broadband	MI	**Logistic Regression**	**87.91 ± 6.34**	**0.893–0.927**

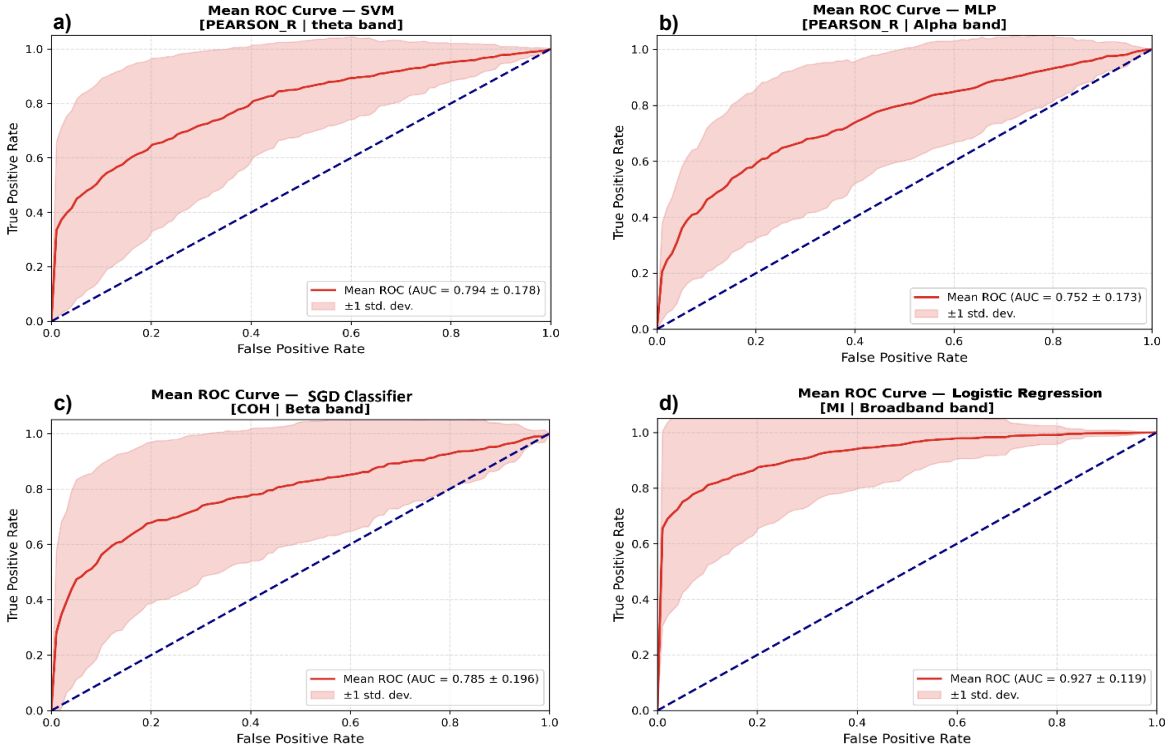

Figure 8. Mean ROC curves for the best-performing LOSO model in each frequency band. **(a)** Theta band SVM with Pearson features (AUC = 0.794 ± 0.178); **(b)** Alpha band—MLP with Pearson features (AUC = 0.752 ± 0.173); **(c)** Beta band—SGD with COH features (AUC = 0.785 ± 0.196); **(d)** Broadband Logistic Regression with MI features (AUC = 0.927 ± 0.119). Shaded regions indicate ± SD across LOSO folds. Dashed blue line = chance level.

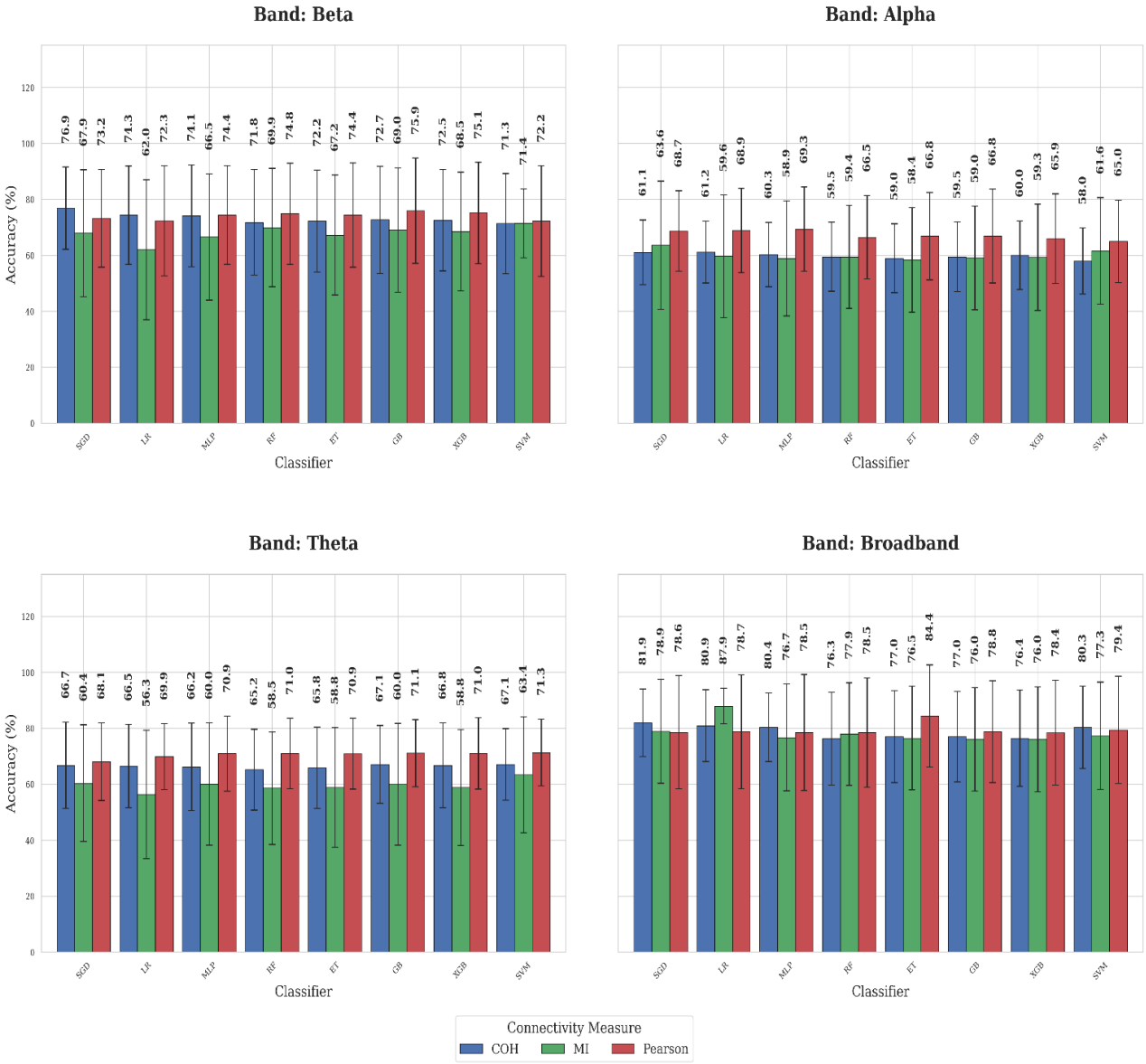

Figure 9. Faceted grouped bar chart showing LOSO classification accuracies for all classifiers (Logistic Regression, SGD, MLP, Random Forest, Extra Trees, Gradient Boosting, XGBoost, and SVM) across connectivity types (COH, Pearson correlation, MI) and frequency bands (theta, alpha, beta, broadband).

In summary our results demonstrate that EEG-based attention classification depends on both the choice of connectivity measure and frequency band. Subject-dependent models consistently achieved higher accuracies than LOSO, reflecting the advantage of within-subject optimization, with MI features combined with Logistic Regression yielding the highest performance across all bands. LOSO classification indicated that broadband MI features tended to outperform single-band analyses and linear measures; however, the relatively high standard deviation across folds highlights variability and instability in cross-subject predictions within this small sample. Finally, comparing our findings with existing literature (Table 4) suggests that graph-based connectivity measures offer superior discriminative power compared to conventional single-electrode spectral or entropy-based features.

Table 4. Classification performance comparison with related works.

Related Work	Feature Type	Evaluation Criteria and Accuracy
Acı *et al.* [23]	STFT (using Blackman window)	Subject dependent 5-fold cross validation Accuracy = 91.72% Subject independent Accuracy = 78.8%
Christina *et al.* [24]	ERP, time-frequency analysis	LOSO Accuracy = 50%–85%
Peng *et al.* [25]	Hilbert-Huang Transform (HHT)	Subject Independent Accuracy = 83.90%
Enrique *et al.* [7]	Time-domain features	Subject dependent Accuracy = 88.9%
Vortmann *et al.* [8]	PSD features	Subject dependent 5-fold cross-validation Accuracy = 85.37% ± 11.27%
Fan *et al.* [9]	Entropy-based features	Subject dependent 5-fold cross-validation Accuracy = 92.31%
Our Work	Graph Based Features	Subject dependent 3-fold cross-validation Accuracy = **97.23%** Subject Independent Accuracy = **87.91%**

4. Discussion

This study examined graph-theoretical features derived from three complementary functional connectivity measures Coherence, Pearson correlation, and Mutual Information—across theta, alpha, and beta frequency bands to distinguish sustained attentional engagement from passive viewing in a ten-participant preliminary investigation. Our results results suggest that attentional states are associated with dissociable network topologies that differ across frequency bands, and that linear and nonlinear connectivity measures capture partially complementary aspects of these network differences. We emphasize throughout that all neurophysiological observations are preliminary given the small sample size and require independent replication before they can be considered established effects.

4.1. Frequency-band network signatures

Our theta-band findings tentatively suggest a small-world-like network organization during sustained attention, characterized by higher node strength, higher clustering coefficient, and shorter average path length compared to inattention [26].This pattern higher local clustering combined with shorter global path lengths is consistent with the theoretical framework proposing that small-world topology supports optimal cognitive processing through a balance of local specialization and global integration [27]. The dominance of COH and Pearson measures in capturing theta-band differences, with no MI features surviving FDR correction, tentatively suggests that linear oscillatory synchronization is the primary coupling mechanism underlying theta-band attentional modulation in this dataset. This aligns with the role of frontal-midline theta as a mechanism for coordinating distributed cognitive resources through phase synchrony [28], and with prior graph-theoretical work linking theta-band network integration to working memory performance [27]. However, the difference between working memory tasks in that prior work and the sustained naturalistic attention paradigm used here limits the strength of this comparison, and replication with controlled laboratory paradigms is needed.

In the alpha band, we observed a contrasting pattern: during inattention, higher betweenness centrality and longer, more variable path lengths indicate a hub-dominated, segregated topology, while

the attention condition appears associated with more locally integrated network organization. The convergence of COH, MI, and Pearson measures on longer alpha path length variability during inattention provides some triangulation across methodologically distinct coupling frameworks. This pattern is broadly consistent with alpha oscillations' proposed role in active cortical inhibition during passive or internally directed states [29], though we emphasize that these are preliminary observations from a small sample and independent replication is essential before these network patterns can be considered reliable markers of attentional modulation in the alpha band.

Beta-band results yielded the richest pattern of significant features, suggesting that sustained attention is associated with more integrated, densely coupled network topology higher global efficiency, higher node strength, higher clustering coefficient, and shorter path lengths while passive viewing is associated with more segregated network organization. The convergence of COH and MI on longer beta path lengths during inattention provides cross-measure consistency and adds modest plausibility to these preliminary observations. The Pearson-derived global efficiency finding higher during attention is tentatively consistent with beta oscillations' proposed role in maintaining task-relevant sensorimotor and cognitive states [30], and with prior reports associating higher beta-band global efficiency with better cognitive performance [31]. However, the basis of this finding in a single connectivity measure in a small sample warrants caution, and it should be treated as exploratory. Across all three frequency bands, path length emerged as the most consistently significant feature, surviving FDR correction in all bands and across multiple connectivity types, suggesting it may be a particularly sensitive marker of attentional network organization in this dataset.

4.2. Complementarity of connectivity measures

A key motivation of this study was to evaluate whether linear and nonlinear connectivity measures capture distinct aspects of attentional network dynamics. Our results provide some preliminary support for this premise: theta-band differences were captured exclusively by linear measures (COH and Pearson), while MI contributed uniquely to alpha-band path length variability and to beta-band path length detection alongside COH. MI-based broadband features also provided the most stable cross-subject LOSO classification performance, with the lowest standard deviation across folds. This pattern is compatible with the idea that nonlinear coupling captured by MI reflects aspects of neural interaction that linear measures do not fully characterize [16], though the functional significance of this difference in the context of attentional modulation remains to be established. It is also possible that the relative sensitivity of MI versus linear measures reflects differences in signal-to-noise ratio or estimation variance rather than genuine differences in the underlying neural coupling mechanisms, and this alternative cannot be excluded with the current data.

4.3. Classification Performance and Generalizability

Subject-dependent classification substantially outperformed LOSO across all bands and connectivity types, which is expected given the absence of subject-specific calibration in LOSO and the small training set in each fold (N = 9 subjects). The 97.23% subject-dependent accuracy achieved by MI features with Logistic Regression in the beta band should be interpreted cautiously: with only 10 participants and 3-fold cross-validation, this estimate carries substantial variance, and the true population accuracy is likely

lower. More informative are the LOSO results, which reflect the more challenging and practically relevant scenario of cross-subject generalization. The highest LOSO accuracy was achieved by broadband MI features with Logistic Regression (87.91% ± 6.34%), which also showed the lowest standard deviation across folds, suggesting more stable generalization than single-band models. This is consistent with the broader principle that multi-band integration captures more information than any single frequency band [22], and with MI's ability to capture nonlinear dependencies that may generalize more consistently across subjects than traditional frequency- or entropy-based features.

Comparison with related work (Table 4) highlights the advantage of our graph-based connectivity approach over conventional single-channel methods. Subject-dependent accuracy of 97.23% exceeds the performance reported in studies using purely frequency-domain features (Acı *et al.* [23], 91.72%) or entropy-based measures (Fan *et al.* [9], 92.31%). The LOSO accuracy of 87.91% also surpasses traditional frequency- or time-domain feature approaches such as Christina *et al.* [24] (50%–85%) and Peng *et al.* [7] (83.90%), demonstrating that graph-theoretical features derived from multichannel interactions capture more discriminative information for attentional state detection than features extracted from individual channels. Notably, our results were obtained with a deliberately low-density eight-channel configuration, further emphasizing the efficiency and robustness of the proposed connectivity-based framework compared to classical single-channel analyses.

4.4. Limitations

Our study faces several constraints that warrant discussion, starting with our modest participant cohort. Statistical power is limited, and both p-values and effect sizes should be interpreted as preliminary indicators rather than established characterizations. The reported effect sizes, while large, may be inflated due to small-sample estimation bias [19], and the FDR correction applied here controls Type I error but does not address the reduced power for detecting true effects. The absence of behavioural ground truth such as comprehension testing, eye tracking, or reaction time measures means that condition labels are based entirely on task assumption rather than verified individual attentional states, which limits the interpretive strength of all network-level observations. The use of a naturalistic paradigm, while ecologically motivated, introduces the possibility of differences in sensory processing, cognitive fatigue, and motivation that are not exclusively attentional, and these confounds cannot be ruled out with the current design. Additionally, the single experimental session per participant precludes assessment of test-retest reliability of the observed network patterns. Future work should expand the sample size to at least 30 participants, incorporate behavioural verification of attentional states, use counterbalanced within-subject designs to control for order effects and fatigue, and apply the proposed framework to independent validation cohorts before drawing conclusions about its generalizability for real-world cognitive monitoring applications.

5. Conclusion

Our study investigated band wise graph-theoretical features derived from multichannel EEG connectivity as a framework for distinguishing sustained attentional engagement from passive inattention. Across all three frequency bands, preliminary network-level differences were observed after Benjamini-Hochberg FDR correction, with path length emerging as the most consistently significant

feature across bands and connectivity types. Theta-band networks tentatively showed small-world-like topology during attention; alpha-band networks showed more hub-dominated, segregated topology during inattention; and beta-band networks showed the richest pattern of significant features, including higher global efficiency, stronger coupling, and more integrated topology during attention.

Classification analyses demonstrated that MI-based features combined with Logistic Regression achieved the highest subject-dependent accuracy (97.23% in the beta band) and the most stable cross-subject LOSO performance (87.91% ± 6.34% for broadband MI), with the notably low standard deviation of the broadband LOSO result indicating genuine cross-subject consistency. These results tentatively support the complementary use of linear oscillatory, linear amplitude, and information-theoretic connectivity measures to characterize attentional network dynamics, and suggest that broadband MI features may provide a robust basis for generalizable EEG-based attention detection.

These findings should be interpreted as preliminary evidence motivating larger-scale replication rather than definitive characterizations of attentional network topology. The small sample size, reliance on task-assumption-based condition labeling, and single-session design substantially limit the generalizability of the reported observations. Future work incorporating larger samples, behavioural ground truth, and independent validation cohorts is needed to establish whether these network-level markers reliably distinguish attentional states in ecologically valid settings.

Data availability statement

The data supporting the findings of this study are not publicly available due to participant privacy and the fact that the dataset is currently being utilized in ongoing, separate research projects. However, the data are available from the corresponding author upon reasonable request.

Declaration of generative AI and AI-assisted technologies

We would like to state that tools such as Grammarly and Claude were used only for language polishing, grammar correction, and improving the clarity and readability of the text. The authors take full responsibility for the content of the manuscript.

Acknowledgments

This work was funded by the Lebanese university, program of innovation and development.

Authors' contribution

Sami Al Majanini: conceptualization, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, visualization; Mohamad Khalil: conceptualization, validation, resources, review and editing, supervision, funding acquisition; Wassim El Falou: validation, review and editing, supervision, project administration. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

The researchers report no financial, professional, or personal conflicts of interest that could be construed as influencing the results or interpretation of this study.

References

[1] Li Y, Li X, Ratcliffe M, Liu L, Qi Y, *et al*. A real-time EEG-based BCI system for attention recognition in ubiquitous environment. In *Proceedings of 2011 ACM Workshop on Ubiquitous Affective Awareness and Intelligent Interaction (UAAII'11)*, Beijing, China, September 18, pp. 33–40.

[2] Wan Z, He J, Voisine A. An attention level monitoring and alarming system for the driver fatigue in the pervasive environment. In *Proceedings of International Conference on Brain and Health Informatics (BHI 2013)*, Maebashi, Japan, October 29–31, 2013, pp. 287–296.

[3] Arns M, Conners C, Kraemer H. A decade of EEG theta/beta ratio research in ADHD. *J. Atten. Disord.* 2013, 17(5):374–383.

[4] Värbu K, Muhammad N, Muhammad Y. Past, present, and future of EEG-based BCI applications. *Sensors* 2022, 22(9):3331.

[5] Fan J, Byrne J, Worden MS, Guise KG, McCandliss BD, *et al*. The relation of brain oscillations to attentional networks. *J. Neurosci.* 2007, 27(23):6197–6206.

[6] Zhou Y, Huang S, Xu Z, Wang P, Wu X, *et al*. Cognitive workload recognition using EEG signals and machine learning: a review. *IEEE Trans. Cognit. Dev. Syst.* 2021, 14(3):799–818.

[7] Martínez Beltrán ET, Quiles Pérez M, López Bernal S, Martínez Pérez G, Huertas Celdrán A. SAFECAR: a brain–computer interface and intelligent framework to detect drivers' distractions. *Expert Syst. Appl.* 2022, 203:117402.

[8] Vortmann LM, Kroll F, Putze F. EEG-based classification of internally- and externally-directed attention in an augmented reality paradigm. *Front. Hum. Neurosci.* 2019, 13:348.

[9] Fan C, Peng Y, Peng S, Zhang H, Wu Y, *et al*. Detection of train driver fatigue and distraction based on forehead EEG: a time-series ensemble learning method. *IEEE Trans. Intell. Transp. Syst.* 2021, 23(8):13559–13569.

[10] Fiebelkorn IC, Kastner S. Functional specialization in the attention network. *Annu. Rev. Psychol.* 2020, 71(1):221–249.

[11] Sun Q, Zhou Y, Gong P, Zhang D. Attention detection using EEG signals and machine learning: a review. *Mach. Intell. Res.* 2025, 22(2):219–238.

[12] Lee K, Choi KM, Park S, Lee SH, Im CH. Selection of the optimal channel configuration for implementing wearable EEG devices for the diagnosis of mild cognitive impairment. *Alzheimers Res. Ther.* 2022, 14(1):170.

[13] Salmi R, Lu G, Chen Q. SHAP-AAD: DeepSHAP-guided channel reduction for EEG auditory attention detection. *arXiv* 2025, arXiv:2507.03814.

[14] Lindquist MA, Caffo B, Kang J, Simpson S, Pinto M, *et al*. Statistics, data science, and the connectome. *Stat. Data Sci. Imaging* 2025, 2(1):2569894.

[15] Srinivasan R, Winter WR, Ding J, Nunez PL. EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics. *J. Neurosci. Methods* 2007, 166(1):41–52.

[16] Martínez Vásquez DA, Posada-Quintero HF, Rivera Pinzón DM. Mutual information between EDA and EEG in multiple cognitive tasks and sleep deprivation conditions. *Behav. Sci.* 2023, 13(9):707.
[17] Garrison KA, Scheinost D, Finn ES, Shen X, Constable RT. The (in)stability of functional brain network measures across thresholds. *Neuroimage* 2015, 118:651–661.
[18] Zalesky A, Fornito A, Bullmore ET. Network-based statistic: Identifying differences in brain networks. *NeuroImage* 2010, 53(4):1197–1207.
[19] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 1995, 57(1):289–300.
[20] Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 2013, 4:62627.
[21] Stam CJ. Modern network science of neurological disorders. *Nat. Rev. Neurosci.* 2014, 15(10):683–695.
[22] Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 2016, 374(2065):20150202.
[23] Ismail L, Karwowski W. A Graph theory-based modeling of functional brain connectivity based on EEG: a systematic review in the context of neuroergonomics. *IEEE Access* 2020, 8:155103–155135.
[24] Acı Çİ, Kaya M, Mishchenko Y. Distinguishing mental attention states of humans via an EEG-based passive BCI using machine learning methods. *Expert Syst. Appl.* 2019, 134:153–166.
[25] Jin C, Borst JP, van Vugt MK. Predicting task-general mind-wandering with EEG. *Cogn. Affect. Behav. Neurosci.* 2019, 19(4):1059–1073.
[26] Peng C, Chen Y, Chen C, Chen S, Cagneau B, *et al*. An EEG-based attentiveness recognition system using Hilbert–Huang transform and support vector machine. *J. Med. Biol. Eng.* 2019.
[27] Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998, 393(6684):440–442.
[28] Dai Z, De Souza J, Lim J, Ho PM, Chen Y, *et al*. EEG cortical connectivity analysis of working memory reveals topological reorganization in theta and alpha bands. *Front. Hum. Neurosci.* 2017, 11:237.
[29] Cavanagh JF, Frank MJ. Frontal theta as a mechanism for cognitive control. *Trends Cognit. Sci.* 2014, 18(8):414–421.
[30] Mathewson KE, Lleras A, Beck DM, Fabiani M, Ro T, *et al*. Pulsed out of awareness: EEG alpha oscillations represent a pulsed-inhibition of ongoing cortical processing. *Front. Psychol.* 2011, 2:99.
[31] Engel AK, Fries P. Beta-band oscillations—signalling the status quo? *Curr. Opin. Neurobiol.* 2010, 20(2):156–165.