IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Credibility-Adjusted Data-Conscious Clustering Method for Robust EEG Signal Analysis

Fatemeh Divan[1], Teh Ying Wah[1], Kheng Seang Lim[2], and Ali Seyed Shirkhorshidi[3]
[1]Department of Information Systems, University of Malaya, KL, 50603, Malaysia
[2]Department of Medicine, University of Malaya, KL, 50603, Malaysia
[3]Entefy Inc., Palo Alto, 94306, CA, USA

Corresponding author: Teh Ying Wah (e-mail: tehyw@ um.edu.my).

ABSTRACT Clustering neurodata, including electroencephalography (EEG) signals, is crucial for brain-computer interface (BCI) and neurological analysis. However, traditional methods struggle with noise, overlapping distributions, and high-dimensional data. This study presents the Credibility-Adjusted Data Conscious Clustering Method (CADCCM), an adaptive computational learning model for clustering neurodata, including EEG signals. CADCCM improves clustering robustness by dynamically adjusting assignments through credibility updates and optimizing parameters for better accuracy and stability in noisy, high-dimensional data. CADCCM dynamically adjusts cluster assignments by integrating alpha and beta parameters to balance fuzzy membership and credibility. A grid search framework optimizes clustering parameters, and preprocessing techniques (Fourier Transform, Wavelet Transform, and Gaussian filtering) improve feature separability. The method is benchmarked against traditional and recent Clustering methods across 13 datasets. CADCCM achieves superior clustering performance, consistently outperforming baseline methods in Rand Index (RI), F-score, and Cohen's Kappa, particularly in noisy datasets. Gaussian filtering further enhances clustering accuracy. GPU acceleration ensures computational feasibility for large-scale neurodata applications. Additionally, CADCCM outperforms other methods by satisfying all clustering properties, including scale invariance, richness, consistency, and order independence. CADCCM bridges the gap between traditional and advanced clustering by introducing credibility-based updates and optimized parameter selection, leading to improved accuracy, robustness, and efficiency. The method holds promise for applications in cognitive state monitoring, neurological disorder detection, and BCI systems. CADCCM enhances the interpretability, reliability, and scalability of clustering. Future research will focus on real-time clustering, adaptive hyperparameter tuning, and deep-learning-based feature extraction.

INDEX TERMS Brain-Computer Interfaces, Machine Learning, Pattern Clustering, Pattern Recognition, Signal Processing.

I. INTRODUCTION

Neurodata, including electroencephalography (EEG), is crucial for analyzing brain activity and significantly contributes to applications such as brain-computer interfaces (BCIs) and neurological diagnostics. However, clustering such complex, high-dimensional data is challenging due to its inherent noise and non-stationarity. Efficient clustering is critical for real-time decision-making, especially in clinical settings, but traditional methods often struggle with these complexities. Standard clustering algorithms, such as K-means and Fuzzy C-means(FCM), are widely used for EEG analysis but exhibit limitations in handling noise, outliers, and the complex variability inherent in EEG signals. Recent approaches, such as multi-weighted EEG clustering (mwcEEGc) [1], attempt to address these issues but often lack flexibility in accommodating data credibility and dynamic cluster structures.

To overcome these limitations, we propose CADCCM, an adaptive clustering algorithm for neurodata. It integrates credibility-based updates to dynamically adjust cluster assignments, enhancing robustness and improving performance in noisy, high-dimensional data. CADCCM incorporates α and β parameters to dynamically change the influence of data points based on their reliability, thereby improving cluster assignment in noisy and uncertain clinical environments. Furthermore, CADCCM employs a systematic grid search optimization to fine-tune its clustering parameters, ensuring improved accuracy and stability compared to existing methods.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3581598

CADCCM holds significant potential for real-time clinical EEG analysis and patient monitoring by clustering noisy EEG signals and identifying underlying brain patterns, which aids in the diagnosis and monitoring of neurological conditions. Its ability to detect subtle patterns in brain activity makes it especially useful for detecting neurodegenerative disorders by offering early indicators for conditions such as Alzheimer's and Parkinson's. In seizure prediction, CADCCM excels in noisy datasets, enabling the identification of pre-seizure brain states and offering a timely intervention tool for clinical settings. Additionally, CADCCM is well-suited for BCI-based rehabilitation, where real-time adaptive clustering can be used to monitor brain activity and support the development of personalized, interactive therapies for neurological patients. Furthermore, CADCCM's robustness in clustering complex EEG data makes it an ideal tool for cognitive load assessment, enabling feedback in applications such as education, occupational health, and rehabilitation.

This study evaluates CADCCM on publicly available EEG datasets and compares its performance against recent advanced clustering techniques. The results assessed using external evaluation metrics, including evaluation criteria like Rand Index (RI), F-score, and Cohen's Kappa coefficient, demonstrate that CADCCM achieves superior clustering accuracy, particularly in noisy and high-variability EEG datasets. The key contributions of this work are:

1. CADCCM, an adaptive clustering model for neurodata, which adjusts cluster assignments based on credibility-aware updates.
2. A systematic parameter optimization using grid search to fine-tune the clustering parameters and ensure robustness and scalability across noisy, high-dimensional neurodata.
3. A comprehensive evaluation on neurodata (including EEG datasets), demonstrating CADCCM's effectiveness in handling complex and uncertain data structures, proving its suitability for real-time EEG assessment and clinical diagnostics.

The structure of this paper is as follows: Section 2 presents the literature background, discussing traditional and recent clustering techniques applied to neurodata analysis, along with their strengths and limitations. Section 3 details the proposed CADCCM framework, including its membership update mechanism, credibility adjustment, and parameter optimization. It also outlines the experimental setup, describing the datasets, preprocessing techniques, and evaluation metrics. Section 4 presents the evaluation findings, comparing CADCCM with contemporary clustering algorithms and examining the impact of different preprocessing methods on performance. Finally, Section 5 concludes the study with key findings and future research directions.

II. RELATED WORKS

Clustering is essential for analyzing EEG signals, enabling the identification of intrinsic patterns that support real-time applications. Various unsupervised clustering algorithms have been explored for EEG data analysis, each offering distinct advantages and facing specific limitations. This section reviews primary clustering methodologies, recent advances, and artifact removal techniques in EEG clustering.

A. UNSUPERVISED CLUSTERING OF EEG SIGNALS

Unsupervised methods are widely employed for EEG clustering, particularly in scenarios where labeled data is unavailable. These methods aim to uncover intrinsic patterns in EEG time-series data for applications such as seizure detection, sleep stage classification, and emotion recognition. Table I provides a comparative analysis of the strengths and weaknesses of traditional clustering methods.

1) K-MEANS CLUSTERING

K-means is commonly used owing to its simplicity and efficiency, particularly in neuropsychological studies related to epilepsy, stress, depression, and Alzheimer's disease [2], [3], [4]. However, its sensitivity to how initial cluster centers are selected and the predefined number of clusters leads to inconsistent performance, particularly when dealing with EEG signals that exhibit complex, non-stationary patterns. Hybrid approaches combining K-means with optimization techniques have been proposed to enhance their robustness [5].

2) FUZZY C-MEANS (FCM)

FCM builds upon the K-means algorithm by assigning membership values to data points, allowing for soft clustering, which is particularly useful in handling noise and uncertainty in EEG signals. However, it remains reliant on predefined cluster centers, which can be problematic for high-dimensional EEG data. To enhance its performance, FCM has been integrated with supervised learning methods [6], [7], [8], dimensionality reduction strategies like Principal Component Analysis (PCA) [9], advanced feature extraction approaches like recent Leveraging Large Language Models (LLMs) [10], and optimization algorithms [11], [12]. These hybrid approaches enhance decision boundaries, improve the clustering of ambiguous samples, reduce feature dimensionality, leverage deep semantic representations from LLMs, and optimize cluster assignments.

3) SPECTRAL CLUSTERING

Spectral clustering has demonstrated effectiveness in EEG analysis, particularly for segmenting brain states in neurocritical care patients. By utilizing covariance matrices from short EEG windows, [13] successfully classified burst and suppression phases in burst suppression pattern detection. The algorithm was evaluated using 29 hours of EEG recordings collected from 29 individuals and achieved promising results, reporting an average absolute error of 0.93 bursts per minute. [14] explored spectral clustering for classifying the epileptic EEG dataset. It applied graph Laplacian-based clustering and demonstrated that spectral clustering, with proper preprocessing, outperformed traditional methods, such as k-means and k-medoids, in EEG classification.

4) K-MEDOID

[15]applies K-Medoid to cluster EEG data, aiming to enhance classification accuracy by reducing noise and refining decision boundaries. [16] utilizes K-Medoid for partitioning brain signal patterns, particularly in detecting cognitive states, benefiting from its robustness to outliers. [17] employs Partitioning Around Medoids (PAM), a classic K-Medoid method, to optimize clustering performance in EEG feature space, ensuring stable cluster formation. These studies leverage K-Medoid's ability to provide more representative cluster centers compared to centroid-based methods, such as K-Means, making it well-suited for EEG data analysis.

TABLE I
PROS AND CONS OF TRADITIONAL CLUSTERING METHODS FOR EEG SIGNAL ANALYSIS

Methods	Advantages	Disadvantages
K-means	Simple, fast, and computationally efficient for distinct clusters.	Sensitive to noise, initial centroids, and the number of clusters.
FCM	Flexible membership. Handles uncertain and overlapping EEG data. Avoids local optima.	Sensitive to noise. Sensitive to initialization. Unstable in high dimensions. Risk of overfitting. Requires extra tuning.
Spectral	Captures complex structures and non-linear separability. Offers flexibility.	Computationally expensive for large datasets. Requires careful parameter tuning.
K-Medoid	Offers flexibility. Robust to outliers. Produces stable clusters, less affected by noise. Interpretable and suitable for small to medium datasets.	Computationally expensive. Slow due to pairwise distance calculations. Not scalable for high-dimensional data. Requires parameter tuning. Overlapping clusters.

B. RECENT ADVANCES IN NEURODATA CLUSTERING

With the increasing complexity of Neurodata, recent clustering methods have incorporated advanced techniques such as graph-based clustering, deep learning, and adaptive clustering strategies. [18] employed dimensionality reduction using t-distributed Stochastic Neighbor Embedding (t-SNE) for reducing data dimensionality, followed by clustering through Density-Based Spatial Clustering of Applications with Noise (DBSCAN). This integrated approach reached a classification accuracy of 91% in classifying sleep states and has been beneficial for automated behavioral-state scoring. Lee and Alamaniotis [19] introduced a DESOM (Deep Embedded Self-Organizing Map) model for unsupervised prediction of cybersickness from EEG signals. The model integrates an EEGNet-based autoencoder with a SOM layer to learn feature representations and cluster structures jointly. DESOM outperformed baseline SOM and KNN approaches in purity and NMI, highlighting its effectiveness for clustering high-dimensional neurodata. [20] developed a method for detecting high-frequency oscillations (HFOs) in intracranial EEG (iEEG) to aid seizure onset zone localization. Their approach combines Singular Value Decomposition (SVD) to extract features using an enhanced FCM, optimized using Simulated Annealing and Genetic Algorithms to overcome initialization sensitivity. The approach outperformed traditional methods, highlighting the strength of fuzzy clustering in handling EEG uncertainty. Lekova and Chavdarov [21] developed BCIFS, a fuzzy system for interpretable EEG-based BCI design. Using fuzzy rules and the Sugeno model, this approach models spatiotemporal EEG dynamics, enabling real-time artifact detection and brain-state classification, which is validated through attention decoding in maze navigation tasks.[1] is constructed as a graph where EEG signals are clustered based on their similarity in a Fr´echet distance-weighted graph. This method effectively addresses the challenge of segmenting noisy and complex EEG data, enhancing accuracy and stability; however, it is computationally expensive and sensitive to parameter selection. Adaptive density peak clustering (ADPC) [22] was an unsupervised clustering method designed for EEG-based epilepsy detection. It automatically selected optimal cutoff distances and cluster centers, outperforming traditional methods like K-means, DBSCAN, and density peaks clustering with k-nearest neighbors (DPC-KNN). Zhang [23] introduced the Time-Aware Sampling Network (TAS-Net), designed to detect EEG segments that adapt to dynamic emotional states. This approach enhances the accuracy of EEG-based emotion recognition compared to traditional unsupervised methods. Adaptive Mixture Independent Component Analysis (AMICA) [24] was employed to analyze emotional imagination using EEG recordings. It reveals spatiotemporal EEG patterns associated with different emotions, demonstrating strong potential for exploring brain dynamics. Recent research [25] has proposed deep learning approaches such as Variational Autoencoders (VAEs) for unsupervised seizure detection. This method models standard EEG patterns and identifies seizure-related deviations without requiring labeled data. Unlike traditional methods that focus on resting-state EEG, the Riemannian distance-based approach [26] combined spatial patterns and a deep autoencoder to segment and cluster EEG data during cognitive tasks, aiming to uncover the temporal dynamics of mental processes. Validated using synthetic EEG signals along with data from two cognitive experiments, SPADE outperformed baseline clustering algorithms in accuracy and multiple evaluation metrics. Kawaguchi et al. [27] developed an EEG-based facial gesture recognition system using a Self-Organizing Map (SOM) with Hebbian learning. Instead of removing artifacts, the method utilized muscle-related EEG signals (e.g., blinks, jaw movements) as gesture cues. SOM clustered EEG features, and a Hebbian layer mapped them to gestures, achieving up to 98% accuracy. This emphasizes the promise of unsupervised neural clustering approaches for real-time BCI systems. Zhu et al. [28] introduced SOGPCN, a deep learning model for EEG-based emotion recognition that combines attention, pseudo-3D convolution, and self-organizing graph modules. By modeling spatial and temporal EEG features, the method achieved up to 95.26% accuracy on the SEED dataset, demonstrating the effectiveness of self-organized spatiotemporal learning on neuro-data.

C. UNSUPERVISED ARTIFACTS REMOVAL IN NEURODATA

Eliminating artifacts—like eye blinks, muscle activity, and ambient noise—from EEG recordings is a crucial preprocessing step for accurate analysis. Various unsupervised methods have been introduced to autonomously identify and eliminate artifacts in EEG recordings.

1) WAVELET TRANSFORM-BASED METHODS

Wavelet-based approaches decompose EEG signals into their frequency components, facilitating noise separation. These methods have been widely used due to their effectiveness in processing non-stationary signals. [29] used wavelet transform to generate scalograms and computed three similarity levels (latent, image, and deep feature map similarity) to identify EEG signal clusters. The approach was validated on public datasets, showing effective clustering, with potential applications in epilepsy diagnosis and multi-label EEG analysis. Murali [30] integrated wavelet analysis with Circular Singular Spectral Analysis (Ci-SSA) to eliminate EOG-related noise from single-channel EEG recordings. This approach began with wavelet-based feature extraction, followed by k-means clustering, to improve EEG signal clarity, which is especially beneficial in real-time processing scenarios.

2) CLUSTERING-BASED ARTIFACT SEGMENTATION

K-means and DBSCAN have been utilized to segment EEG signals into clean and contaminated regions, aiding in automatic artifact removal[31]. Moreover, an integrated thresholding technique utilizing a Gaussian Mixture Model (GMM) has been introduced to identify eye-blink artifacts, demonstrating enhanced detection precision [32].

3) RECENT CLUSTERING METHODS FOR NOISE REMOVAL

A deep learning-based unsupervised outlier detection method [33] has improved artifact correction in EEG data. One approach combined a deep encoder-decoder network with traditional outlier detection techniques to enhance EEG signal quality. Du et al. [34] proposed an unsupervised adaptive clustering algorithm for epileptic EEG signals, leveraging CEEMDAN, continuous wavelet transform (CWT), and t-SNE for feature extraction before applying DBSCAN with an adaptive parameter selection strategy. A multi-set consensus clustering technique has been introduced [35] to identify stable cognitive responses while mitigating noise, aiming to evaluate individual EEG trials and measure event-related brain responses (ERPs). This method improves subject-specific EEG analysis, mitigating noise effects. Despite advancements in EEG clustering, challenges remain as clustering methods often assume simplistic data distributions, are sensitive to noise and outliers, and fail to adapt to dynamic EEG signals. Table II highlights the strengths and weaknesses of these methods.

TABLE II
COMPARISON OF RECENT CLUSTERING METHODS FOR EEG SIGNAL ANALYSIS

Methods	Advantages	Disadvantages
[1]	Accurate, stable with noisy data.	Computationally intensive. Parameter-sensitive. Limited generalizability across datasets.
[18]	Handles Noise Effectively. Does Not Require Predefined Clusters. Handles High-Dimensional Data.	Computationally Intensive. Sensitive to Parameter Selection.
[19]	Combines deep learning with SOM for better clustering. Effective for complex EEG data.	High computational cost. Needs large, labeled datasets for training.
[20]	Handles uncertainty in over-lapping distributions. Effective for fuzzy relationships in EEG	Sensitive to initial cluster centers. Struggles with high-dimensional data.
[21]	Combines fuzzy clustering and SOM for effective EEG analysis. Improves BCI performance.	Complex system needing tuning. High computational resources needed for real-time applications.
[22]	Finds clusters automatically. Handles complex shapes.	High computational load. Parameter sensitivity.
[23]	Adapts to dynamic states in EEG. High emotion recognition accuracy.	High computational cost. Requires careful model and hyperparameter tuning.
[24]	Identifies dynamic brain patterns. Good for emotional analysis.	Computationally intensive. Variable results across subjects.
[25]	No labeled data needed. Detects anomalies well.	High training complexity. Resource-intensive.
[26]	Captures spatial patterns. Accurate for cognitive tasks.	Complex. Requires significant computing power.
[27]	Visualizes high-dimensional data. Preserves relationships between data points.	Requires manual tuning. Struggles with noisy data and outliers.
[28]	Improves computational efficiency. Increases accuracy in emotion/ cognitive state recognition.	May underperform with noisy EEG. Requires careful parameter selection.
[29]	Effective for non-stationary signals. Separates noise well.	Needs careful setup. Can introduce artifacts if improperly configured.
[30]	Removes EOG artifacts without labeled data. Enhances signal quality.	Computationally complex. Performance varies with signals.
[31]	Auto-segments clean/noise data. Real-time potential.	Sensitive to parameters. May misclassify subtle artifacts.
[32]	High detection precision. Robust feature extraction.	Struggles with subtle artifacts.
[33]	Corrects artifacts effectively. Handles complex data.	Needs large datasets. Computationally demanding.

[34]	Improves epileptic EEG clustering. Flexible parameter selection.	Computationally intensive. Needs fine-tuning for datasets
[35]	Improves subject-specific analysis. Reduces noise.	Time-consuming. Computationally expensive.

While existing EEG clustering methods demonstrate promising results, they face key challenges: scalability is an issue, as large EEG datasets require efficient clustering for real-time use. Many unsupervised methods struggle with noise in artifact-heavy environments, while techniques such as K-means and FCM are susceptible to parameter selection, which affects their reliability. Our proposed Credibility-Adjusted Data-Conscious Clustering Method (CADCCM) addresses these limitations by introducing credibility adjustments using alpha and beta parameters to dynamically adjust the influence of each data point based on its reliability. CADCCM further enhances clustering performance through systematic grid search optimization, ensuring robust results even in noisy EEG environments. The integration of credibility adjustment and parameter optimization in CADCCM enhances the reliability of EEG clustering. By weighing data points based on their reliability, CADCCM reduces the impact of noise and artifacts. Grid search optimization fine-tunes parameters, thereby improving the adaptability, complexity, and variability of EEG data. These features make CADCCM more robust than traditional methods. To validate the effectiveness of CADCCM, it has been compared against several existing clustering algorithms, including K-means, FCM, SOM, mwcEEGc, K-medoid, Spectral clustering, and rough set clustering.

III. EXPERIMENTAL SETUP AND METHODOLOGY

A. DATA COLLECTION AND PREPROCESSING

To assess the CADCCM framework, 13 EEG datasets were used, encompassing data types such as Slow Cortical Potentials (SCPs), mental imagery tasks, motor imagery recordings, and hand-movement EEG[5]. Table III demonstrates the details of the EEG dataset. Although originally labeled, the labels were removed before clustering to simulate an unsupervised learning environment.

TABLE III
EEG DATASET CHARACTERISTICS

Data set	Description	EEG Trials (Length)	Channels	Classes
II_Ia	SCPs recorded from a healthy participant.	268 × 5377	6	2
II_Ib	SCPs obtained from an ALS patient.	200 × 8065	7	2
III_V_s1 III_V_s2 III_V_s3	Mental imagery data from three healthy individuals (left/right hand movement & word association).	3488 × 97 3472 × 97 3424 × 97	8	3
IV_2a_s1 IV_2a_s2 IV_2a_s3	Multi-class motor imagery involving three healthy subjects (left/right hand, both feet & tongue).	288 × 6887	22	4
IV_2b_s1 IV_2b_s2 IV_2b_s3	Motor imagery of left/right hand movements from three subjects.	120 × 940	3	2
IV_3_s1 IV_3_s2	Hand movement data from two healthy subjects performing 4 directional tasks.	160 × 4001	10	4

Significant noise was observed upon visualizing the raw EEG signals, which may have affected the clustering accuracy. To address this, noise reduction techniques were applied to enhance signal quality before clustering.

B. PREPROCESSING PIPELINE

Preprocessing ensures that the EEG data is in an optimal format for clustering, enhancing noise reduction, and dimensionality reduction while preserving critical features. This methodology combines preprocessing methods with our proposed CADCCM framework to compare their results and identify the most effective approach for enhancing clustering performance.

1) FOURIER TRANSFORM[36]:

Converts time-domain EEG signals into the frequency domain, highlighting key frequency components associated with cognitive states. The Fourier Transform is defined as:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt \qquad (1)$$

, where *x(t)* is the time-domain signal and *X(f)* represents its frequency-domain components.

2) Z-NORMALIZATION[1]:

Scales the data such that the mean becomes zero and the standard deviation equals one, promoting uniformity across channels. The z-normalization formula is:

$$Z = \frac{x - \mu}{\sigma} \qquad (2)$$

, where x denotes the original value, μ represents the average, and σ indicates the data's standard deviation.

3) WAVELET DENOISING[37], [38]:

The Wavelet Transform breaks down a signal into scaled and translated forms of a base wavelet, enabling detailed examination of both short-term (high-frequency) and long-term (low-frequency) features. The formula for the Discrete Wavelet Transform (DWT) is:

$$w(j,k) = \sum_n x(n).\psi_{j,k}(n) \qquad (3)$$

, where: $w(j,k) =$ Wavelet coefficient at scale j and position k.

- j (Scale parameter): Controls the frequency resolution.

– Small j →High frequencies (detailed features).

– Large j →Low frequencies (coarse features).

- k (Translation parameter): Determines the position of the wavelet along the time axis, providing time localization.
- $\psi(n)$ (Mother wavelet): A prototype function that generates wavelets through scaling and shifting.
- $\psi_{j,k}(n)$ Scaled and translated version of the mother wavelet.

4) GAUSSIAN FILTER[39]:

The Gaussian filter smooths data by convolving the signal with a Gaussian function, effectively reducing high-frequency noise while preserving the overall shape of the signal.

$$G(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-t^2/_{2\sigma^2}} \qquad (4)$$

, where σ controls the spread of the Gaussian kernel, which influences the degree of smoothing.

C. CLUSTERING ALGORITHM: CADCCM FRAMEWORK

The CADCCM approach refines the fuzzy clustering process by incorporating credibility and parameter optimization:

1. *Membership Update Mechanism:* CADCCM initializes membership matrices randomly and iteratively updates them based on credibility-adjusted weighting.

   $$U_{ij}^{(t+1)} = \frac{\alpha U_{ij}^{(t)} + \beta C_{ij}^{(t)}}{\sum_{k=1}^{c} \alpha U_{ik}^{(t)} + \beta C_{ik}^{(t)}} \qquad (5)$$

   , where α controls the influence of standard fuzzy membership, and β adjusts the impact of credibility scores. C_{ij} is the credibility score calculated as:

   $$C_{ij} = \frac{1}{1 + D_{ij}} \qquad (6)$$

   , where D_{ij} denotes the Euclidean distance between the point X_i and the cluster center V_j. The credibility of each data point influences its membership update, with adjustments normalized to maintain consistent clustering behavior and mitigate noise sensitivity.
2. *Convergence Criterion:* Convergence is determined once the difference in the membership matrix across iterations is smaller than a set threshold:

   $$\|U^{t+1} - U^{t}\|_{\mathrm{F}} < \tau \qquad (7)$$

   , where ‖.‖F is the Frobenius norm.
3. *Cluster Center Calculation:* Computed using weighted averages, ensuring stability against noise. The cluster center V_j is updated iteratively using:

   $$V_j = \frac{\sum_{i=1}^{n} (U_{ij})^m x_i}{\sum_{i=1}^{n} (U_{ij})^m} \qquad (8)$$

   , where m is the fuzzification parameter controlling cluster compactness.
4. *Grid Search Optimization:* To optimize CADCCM, grid search is applied over three parameters:

   $$\boldsymbol{\alpha} \in [\mathbf{0}, \mathbf{1}], \boldsymbol{\beta} = \mathbf{1} - \boldsymbol{\alpha}^{\mathbf{2}}, \boldsymbol{\tau} \in [\mathbf{10^{-3}}, \mathbf{10^{-2}}] \qquad (9)$$

The best parameter combination is selected based on clustering performance metrics, including RI, Kappa, and F-score.

D. Evaluation Metrics

CADCCM's performance was evaluated using evaluation criteria, including the Rand Index (RI), Cohen's Kappa, and F-score.

1) RAND INDEX (RI):

The rand index measures clustering accuracy by counting correctly grouped instances, representing the proportion of accurate assignments:

$$\mathrm{RI} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \qquad (10)$$

, where TP indicates true positives, TN stands for true negatives, FP refers to false positives, and FN denotes false negatives.

2) COHEN'S KAPPA:

Evaluates how well predicted clusters align with actual labels, accounting for agreement that might occur by random chance, emphasizing CADCCM's reliability.

$$k = \frac{(p - p_e)}{(1 - p_e)} \qquad (11)$$

, where p represents the actual agreement rate, while $p_e$is the agreement expected if clusters were assigned randomly.

3) F-SCORE:

Integrates precision and recall, providing a comprehensive measure of clustering effectiveness, especially useful when dealing with class imbalance.

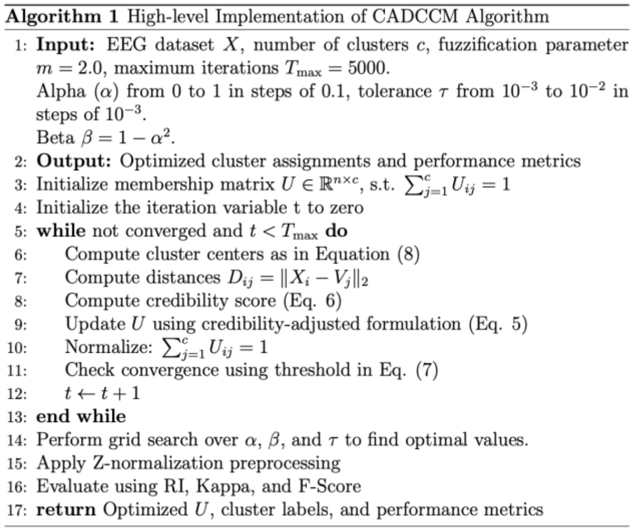
$$\mathrm{F-score} = \frac{(1+\beta^2).p.r}{\beta^2.p + r} \qquad (12)$$

, where $p = \frac{\mathrm{Tp}}{\mathrm{TP+FP}}$ (precision), $r = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$ (recall), and β is commonly set to 1.

External evaluation metrics, including RI, Cohen's Kappa, and F-score, were chosen to measure clustering performance against ground-truth labels available in EEG datasets. Unlike internal metrics, which assess clustering structure without reference labels, external metrics provide a direct and objective measure of clustering accuracy. This ensures a more reliable evaluation of CADCCM's ability to correctly cluster EEG signals, which is critical for biomedical applications.

E. ALGORITHM PSEUDOCODE

A high-level implementation of the CADCCM algorithm is summarized in Fig. 1:



```
Algorithm 1 High-level Implementation of CADCCM Algorithm
Input: EEG dataset $X$, number of clusters $c$, fuzzification parameter $m = 2.0$, maximum iterations $T_{\max} = 5000$.
    Alpha ($\alpha$) from 0 to 1 in steps of 0.1, tolerance $\tau$ from $10^{-3}$ to $10^{-2}$ in steps of $10^{-3}$.
    Beta $\beta = 1 - \alpha^2$.
Output: Optimized cluster assignments and performance metrics
Initialize membership matrix $U \in \mathbb{R}^{n \times c}$, s.t. $\sum_{j=1}^{c} U_{ij} = 1$
Initialize the iteration variable t to zero
while not converged and $t < T_{\max}$ do
    Compute cluster centers as in Equation (8)
    Compute distances $D_{ij} = \|X_i - V_j\|_2$
    Compute credibility score (Eq. 6)
    Update $U$ using credibility-adjusted formulation (Eq. 5)
    Normalize: $\sum_{j=1}^{c} U_{ij} = 1$
    Check convergence using threshold in Eq. (7)
    $t \leftarrow t + 1$
end while
Perform grid search over $\alpha$, $\beta$, and $\tau$ to find optimal values.
Apply Z-normalization preprocessing
Evaluate using RI, Kappa, and F-Score
return Optimized $U$, cluster labels, and performance metrics
```

FIGURE 1. **Pseudocode of the proposed CADCCM method.**

F. EXPERIMENTAL CONFIGURATION

The CADCCM framework is configured with dataset-specific cluster counts, a fuzzification parameter (m) set to

2.0, and a maximum of 5000 iterations or until convergence.

G. TIME COMPLEXITY ANALYSIS OF CADCCM

The computational complexity of CADCCM is influenced by three main factors: (1) credibility-based membership updates, (2) hyperparameter grid search, and (3) evaluation metric computations. In each iteration, CADCCM updates the membership matrix *O(nk)*, computes cluster centers *O(nkd),* and adjusts credibility scores using pairwise distances *O(nkd)*, leading to an overall per-iteration complexity of *O(nkd)*. Given a maximum of *I* iterations, the clustering process runs in *O(Inkd)*. The grid search over alpha(*A*), tolerance values(*T*), and (*R)* runs introduces an overhead of *O(ATRInkd)*. Additionally, clustering performance is evaluated using RI, F-score, and Cohen's Kappa, which operate in *O(n)* complexity. The final computational complexity of CADCCM is *O(ATRInkd)*, which scales efficiently with dataset size and clustering parameters. Unlike mwcEEGc, which relies on Fréchet-based similarity computations with a complexity of

$$O(\max\left\{n^2 l^2\frac{\log\log l}{\log l}, mn^3\right\})$$

[1], CADCCM avoids this overhead, making it more efficient for large-scale EEG clustering while maintaining competitive accuracy.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. PERFORMANCE COMPARISON

CADCCM was benchmarked against recent mwcEEGc[1], SOM, and traditional clustering methods, including: FCM, K-means, K-Medoids, Spectral clustering, and Rough Set Clustering. Experiments were conducted in Python using Colab Notebooks on an iMac with a NVIDIA A100-SXM4-40GB GPU. Each algorithm was run 10 times, and the best performance metrics were reported based on optimal evaluation criteria (either maximum or minimum, depending on the metric). The proposed CADCCM was compared explicitly to mwcEEGc due to their shared objective of addressing EEG time-series clustering challenges, such as noise sensitivity and high-dimensional feature spaces. While mwcEEGc employs a graph-based clique approach, CADCCM introduces a credibility-enhanced fuzzy clustering framework, making it a suitable benchmark for comparative analysis. Since mwcEEGc applied z-normalization as a preprocessing step, we maintained consistency by applying z-normalization across all methods. This ensured a fair and unbiased evaluation of clustering performance. However, discrepancies in dataset selection strategies of the mwcEEGc method could introduce variability in comparative results. The clustering results presented in Tables IV and VI highlight the effectiveness of CADCCM compared to mwcEEGc, SOM, FCM, K-means, K-medoids, Spectral clustering, and Rough Set clustering. The performance evaluation is based on the Rand Index (RI) and F-score, which assess clustering consistency and accuracy, respectively. The mean and variance of the Rand Index and F-score across all datasets are reported in Tables V and VII for a statistical evaluation of the mentioned methods.

TABLE IV
CLUSTERING PERFORMANCE ACROSS EEG DATASETS (RAND INDEX)

Datasets	II_Ia	II_Ib	III_V_s1	III_V_s2	III_V_s3	IV_2a_s1	IV_2a_s2	IV_2a_s3	IV_2b_s1	IV_2b_s2	IV_2b_s3	IV_3_s1	IV_3_s2
RS	0.517	0.497	0.536	0.412	0.446	0.408	0.554	0.582	0.503	0.497	0.496	0.518	0.586
Spectral	0.498	0.501	0.524	0.427	0.383	0.636	0.608	0.594	0.496	0.497	0.496	0.627	0.621
SOM	0.499	0.501	0.559	0.556	0.344	0.64	0.617	0.625	0.501	0.505	0.496	0.623	0.627
FCM	0.498	0.502	0.543	0.547	0.567	0.599	0.635	0.63	0.501	0.505	0.499	0.617	0.579
K-Medoid	0.5	0.5	0.569	0.543	0.555	0.625	0.617	0.626	0.496	0.498	0.496	0.599	0.621
K-means	0.55	0.502	0.598	0.585	0.515	0.645	0.629	0.607	0.502	0.501	0.498	0.627	0.632
CADCCM	0.565	0.526	**0.615**	**0.605**	**0.576**	0.648	**0.637**	**0.642**	0.564	0.57	**0.558**	**0.643**	**0.645**
mwcEEGc	**0.619**	**0.56**	0.602	0.582	0.564	**0.653**	0.635	0.64	**0.583**	**0.575**	0.552	0.637	0.644

TABLE V
MEAN AND VARIANCE OF RAND INDEX VALUES ACROSS EEG DATASETS

Datasets	Mean	variance
RS	0.504	0.003
Spectral	0.531	0.006
SOM	0.546	0.007
FCM	0.556	0.003
K-Medoid	0.557	0.003
K-means	0.569	0.003
CADCCM	0.600	0.002
mwcEEGc	0.604	0.001

TABLE VI
CLUSTERING PERFORMANCE ACROSS EEG DATASETS (F-SCORE)

Datasets	II_Ia	II_Ib	III_V_s1	III_V_s2	III_V_s3	IV_2a_s1	IV_2a_s2	IV_2a_s3	IV_2b_s1	IV_2b_s2	IV_2b_s3	IV_3_s1	IV_3_s2
RS	0.567	0.344	0.323	0.282	0.244	0.191	0.253	0.215	0.547	0.373	0.352	0.24	0.232
K-Medoid	0.535	0.536	0.486	0.381	0.3523	0.31	0.298	0.319	0.5	0.534	0.551	0.305	0.325
Spectral	0.505	0.541	0.43	0.57	0.434	0.345	0.291	0.371	0.509	0.526	0.517	0.312	0.32
SOM	0.524	0.54	0.498	0.408	0.291	0.425	0.292	0.334	0.55	0.569	0.607	0.313	0.325
FCM	0.509	0.545	0.388	0.505	0.448	0.417	0.33	0.364	0.55	0.567	0.533	0.369	0.359
K-means	0.66	0.545	0.556	0.561	0.402	0.393	0.362	0.388	0.56	0.607	0.607	0.377	0.377
mwcEEGc	0.575	0.581	0.563	**0.76**	0.407	**0.475**	**0.465**	**0.463**	0.583	0.568	0.549	**0.435**	**0.43**
CADCCM	**0.683**	**0.62**	**0.6**	0.539	**0.452**	0.419	0.37	0.412	**0.687**	**0.675**	**0.642**	0.407	0.401

TABLE VII
MEAN AND VARIANCE OF F-SCORE VALUES ACROSS EEG DATASETS

Datasets	Mean	variance
RS	0.320	0.014
K-Medoid	0.418	0.011
Spectral	0.436	0.010
SOM	0.437	0.014
FCM	0.453	0.007
K-means	0.492	0.012
mwcEEGc	0.527	0.009
CADCCM	0.531	0.015

Before analyzing the results, it is essential to note that the specific preprocessing steps used in mwcEEGc were not fully detailed in the original publication, making the exact reproduction of its outcomes challenging. Therefore, we report the values as originally presented in their paper (Tables III-XVI) [1]. While this allows us to include mwcEEGc in our comparison, it may lead to an uneven evaluation, particularly in contrast with our proposed CADCCM method, which uses only z-normalization as its preprocessing step. A more equitable comparison would be to compare our method, using preprocessing that optimizes its result, with the mwcEEGc, which we will present in the next section, B.

1) COMPARISON WITH MWCEEGC

This study includes the mwcEEGc method due to its recent development and relevance to advancing EEG clustering methodologies. Table IV shows that the mwcEEGc method achieves competitive RI scores in some datasets, such as II_Ia, II_Ib, and IV_2a_s1. Furthermore, in F-score results (Table VI), CADCCM outperforms mwcEEGc in multiple datasets, including II_Ia, II_Ib, III_V_s1, and IV_2b_s1, which demonstrates its robustness in identifying meaningful clusters, especially in complex and noisy EEG signals. F-score is particularly relevant for EEG clustering as it balances precision and recall, ensuring that clusters are correctly formed and meaningful in distinguishing EEG patterns. The higher F-scores of CADCCM in several datasets indicate better overall clustering quality.

2) COMPARISON WITH OTHER CLUSTERING METHODS

- *K-Means & K-Medoids*: Both traditional centroid-based clustering methods struggle with high-dimensional EEG data and noise sensitivity. Their reliance on predefined centroids often leads to suboptimal partitions, as seen in their lower RI and F-Scores across multiple datasets.
- *FCM*: While FCM allows for soft clustering, it remains sensitive to initialization and noise. CADCCM incorporates credibility adjustments that help mitigate noise sensitivity, improving clustering stability, as reflected in consistently higher F-Scores.
- *Spectral Clustering & Rough Set (RS) Clustering*: These methods are effective in specific scenarios but tend to suffer from high computational costs and sensitivity to parameter tuning. CADCCM, by contrast, strikes an effective trade-off between efficiency and clustering precision by leveraging its credibility-aware design.
- *Self-Organizing Maps (SOMs):* SOMs effectively cluster high-dimensional EEG signals by transforming complex data patterns into lower-dimensional representations. However, SOMs struggle with noise sensitivity, impacting their clustering accuracy. In contrast, CADCCM introduces credibility adjustments that help maintain high clustering accuracy, even in noisy data. This results in CADCCM achieving consistently higher F-scores compared to SOM.

Statistical findings support CADCCM as a robust and reliable method for clustering EEG signals. It consistently delivers high accuracy, with a mean RI of 0.600 and an F-score of 0.527, both of which are among the top performers in their respective categories. Moreover, its low variance values (0.002 for RI and 0.009 for F-score) highlight its stability and reproducibility across diverse EEG datasets. This makes CADCCM particularly suitable for applications where consistency and reliability are essential. The results suggest that CADCCM performs well despite noise in EEG data. Since mwcEEGc likely benefits from undocumented preprocessing techniques, the direct comparison between

CADCCM and mwcEEGc may not be entirely fair. The following section will apply filtering techniques, such as the Fourier Transform, Discrete Wavelet Transform (DWT), and Gaussian filtering, to further enhance CADCCM's performance. However, due to the lack of clarity in mwcEEGc's preprocessing methodology, using the same filters on mwcEEGc would not be methodologically sound, as it could lead to inconsistencies in comparison.

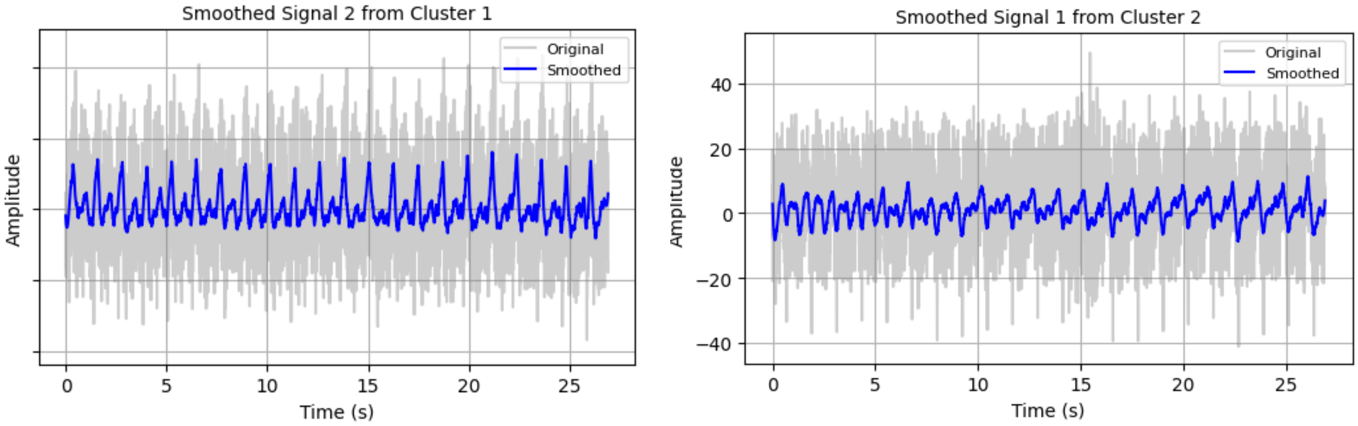
3) DISCUSSION

Overall, CADCCM demonstrates superior performance in multiple datasets, particularly regarding F-score, which indicates a meaningful clustering of EEG analysis. The method's credibility-adjusted framework enables better handling of noisy and overlapping data, distinguishing it from traditional clustering approaches. In the next section, the impact of preprocessing on CADCCM will be further analyzed to optimize its performance. Statistical results demonstrate that CADCCM strikes a strong balance between high clustering performance and robustness, making it highly suitable for applications where consistency and reliability are essential.

B. IMPACT OF PREPROCESSING ON CADCCM PERFORMANCE

To assess the impact of preprocessing on clustering accuracy, we applied the Fourier Transform, Discrete Wavelet Transform (DWT), and Gaussian filtering alongside z-normalization to CADCCM. These preprocessing techniques effectively reduced noise and enhanced signal quality, significantly improving clustering performance, particularly for datasets with high artifact contamination. Fig. 2 visually compares raw vs. denoised EEG signals from the IV_2a_s3 dataset, highlighting four representative signals from each class. From the figure, we can conclude the following:

- *Noise Reduction Effectiveness:* The original EEG signals (gray) contain substantial noise, as indicated by high-frequency fluctuations and random variations. The smoothed signals (blue) exhibit a more transparent structure with reduced noise, showing the effectiveness of the applied preprocessing techniques.
- *Preservation of Underlying Patterns:* Despite the noise reduction, the key waveform structures remain intact in the smoothed signals. This suggests that the applied preprocessing methods (Fourier Transform, Discrete Wavelet Transform, and Gaussian filtering) effectively reduce noise without distorting the essential EEG features.
- *Improved Signal Interpretability:* The cleaned signals exhibit more distinct and regular oscillatory patterns, which can enhance the reliability of clustering by improving feature extraction and reducing noise-induced variability.
- *Potential for Enhanced Clustering Performance*: Since EEG data is susceptible to noise, preprocessing is crucial for improving clustering accuracy. After noise reduction, the more precise and structured signals are expected to result in better-defined clusters, leading to enhanced evaluation metrics.



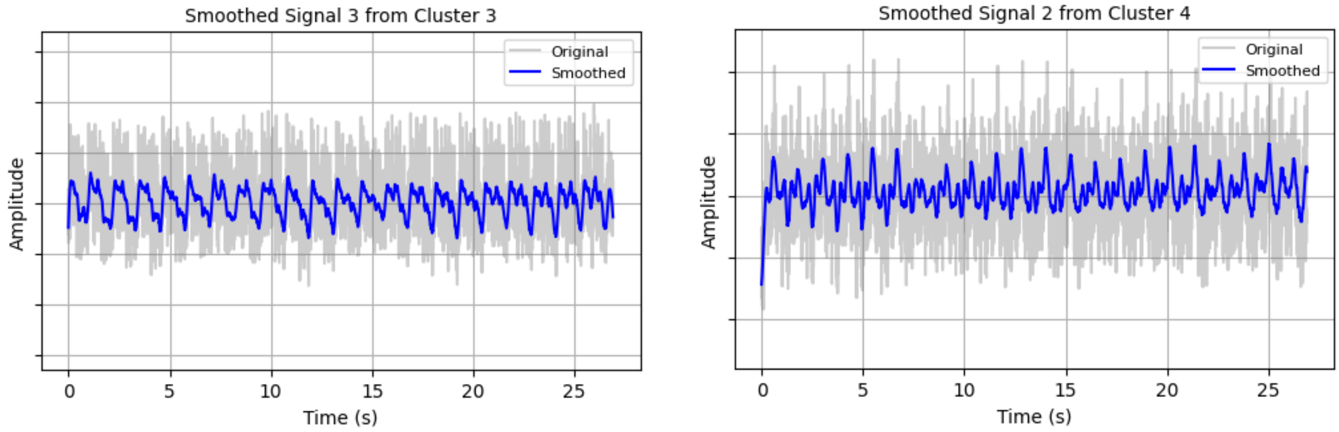


FIGURE 2. **Comparison of raw and smoothed EEG signals after preprocessing.**

Overall, the observed improvements underscore the importance of noise reduction techniques in optimizing EEG clustering, reinforcing CADCCM's robustness in handling real-world EEG signals. Each noise reduction method was first applied to the EEG dataset to denoise signals, followed by z-normalization and clustering with CADCCM, to determine the optimal preprocessing strategy for enhancing clustering accuracy and stability. Tables VIII, IX, and X further demonstrate how various filtering methods influence clustering outcomes, assessed using RI, F-score, and Kappa metrics.

The findings shown in tables VIII, IX, and X demonstrate the impact of various preprocessing techniques (z-normalization, Fourier Transform, Gaussian filtering, and Wavelet Transform) on the clustering performance of CADCCM across multiple EEG datasets. The performance was assessed using three key evaluation metrics: Rand Index (RI), F-score, and Kappa. Gaussian filtering yielded the best average performance across all metrics, with moderate variance, suggesting it enhances EEG signal quality by effectively reducing noise while preserving relevant features for clustering.

1. Rand Index (RI) (Table VIII): Gaussian filtering consistently outperformed other methods, achieving the highest RI scores in datasets such as II_Ia (0.993), IV_2b_s2 (0.935), and IV_2a_s3 (0.921). It yields the highest mean RI (0.839), ahead of Fourier (0.635), Z-norm, and Wavelet (both 0.599). Despite a slightly higher variance (0.010), it provided more reliable and well-separated clustering results.
2. F-score (Table IX): Gaussian filtering achieved the highest F-score in datasets, such as II_Ia (0.996), IV_2b_s1 (0.952), and IV_2b_s2 (0.977), minimizing misclassifications. It also had the highest mean F-score (0.831), with acceptable variance (0.013), outperforming Z-norm, Fourier, and Wavelet methods.
3. Kappa Score (Table X): Gaussian filtering achieved the highest mean Kappa score (0.308) with relatively low variance (0.012), compared to Fourier (0.153), Z-norm (0.113), and Wavelet (0.106). The highest individual Kappa scores were observed in IV_3_s1 (0.484), III_V_s3 (0.452), and III_V_s2 (0.425), indicating strong agreement with the ground truth labels.

TABLE VIII
COMPARISON OF PREPROCESSING METHODS ON CLUSTERING PERFORMANCE (RAND INDEX) WITH MEAN AND VARIANCE

Datasets	Z-norm	Fourier	Gaussian	Wavelet
II_Ia	0.565	0. 789	**0.993**	0.637
II_Ib	0.526	0. 507	**0.609**	0.512
III_V_s1	0.615	0. 677	**0. 742**	0.631
III_V_s2	0.605	0.665	**0.750**	0.595
III_V_s3	0.576	0.682	**0.817**	0.593
IV_2a_s1	0.648	0.702	**0.794**	0.670
IV_2a_s2	0.637	0.650	**0.830**	0.633
IV_2a_s3	0.640	0.658	**0.921**	0.642
IV_2b_s1	0.564	0.541	**0.904**	0.527
IV_2b_s2	0.570	0.598	**0.935**	0.546
IV_2b_s3	0.552	0.520	**0.875**	0.523
IV_3_s1	0.643	0.670	**0.823**	0.643
IV_3_s2	0.646	0.659	**0.822**	0.637
Mean	0.599	0.635	0.839	0.599
Variance	0.002	0.004	0.010	0.003

TABLE IX
COMPARISON OF PREPROCESSING METHODS ON CLUSTERING PERFORMANCE (F-SCORE) WITH MEAN AND VARIANCE

Datasets	Z-norm	Fourier	Gaussian	Wavelet
II_Ia	0.682	0. 880	**0.996**	0.763
II_Ib	0.620	0. 571	**0.737**	0.587
III_V_s1	0.600	0.685	**0.742**	0.611
III_V_s2	0.539	0.623	**0.770**	0.497
III_V_s3	0.452	0.669	**0.842**	0.516
IV_2a_s1	0.419	0.561	**0.675**	0.484
IV_2a_s2	0.370	0.435	**0.769**	0.372
IV_2a_s3	0.410	0.468	**0.920**	0.365
IV_2b_s1	0.687	0.650	**0.952**	0.625
IV_2b_s2	0.675	0.725	**0.977**	0.659
IV_2b_s3	0.642	0.608	**0.937**	0.622
IV_3_s1	0.407	0.458	**0.803**	0.407
IV_3_s2	0.401	0.413	**0.677**	0.370
Mean	0.531	0.572	0.831	0.529
Variance	0.015	0.012	0.013	0.016

TABLE X

COMPARISON OF PREPROCESSING METHODS ON CLUSTERING PERFORMANCE (KAPPA) WITH MEAN AND VARIANCE

Datasets	Z-norm	Fourier	Gaussian	Wavelet
II_Ia	0.110	0. 2575	**0.332**	0.132
II_Ib	0.082	0.045	**0.127**	0.053
III_V_s1	0.228	0.398	**0.127**	0.293
III_V_s2	0.201	0.353	**0. 425**	0.078
III_V_s3	0.066	0.064	**0.452**	0.123
IV_2a_s1	0.144	0.114	**0.315**	0.129
IV_2a_s2	0.067	0.065	**0.296**	0.066
IV_2a_s3	0.075	0.156	**0.220**	0.077
IV_2b_s1	0.117	0.102	**0.306**	0.083
IV_2b_s2	0.100	0.152	**0.333**	0.087
IV_2b_s3	0.074	0.049	**0.302**	0.077
IV_3_s1	0.105	0.240	**0.484**	0.090
IV_3_s2	0.106	0.101	**0.403**	0.090
Mean	0.113	0.153	0.308	0.106
Variance	0.003	0.014	0.012	0.004

The improved performance of Gaussian filtering is due to the following factors:

- *Effective Noise Reduction:* Unlike z-normalization, which standardizes data but does not explicitly remove high-frequency noise, Gaussian filtering smooths the signals while retaining essential information.
- *Preservation of Important EEG Features:* Fourier and Wavelet Transforms may alter the original signal structure, potentially leading to loss of crucial temporal information. Gaussian filtering, however, balances noise suppression and feature preservation.
- *Better Cluster Separability:* The consistently high RI and F-score values indicate that Gaussian filtering enhances the clarity of EEG data, resulting in more distinct clusters that improve overall clustering reliability.

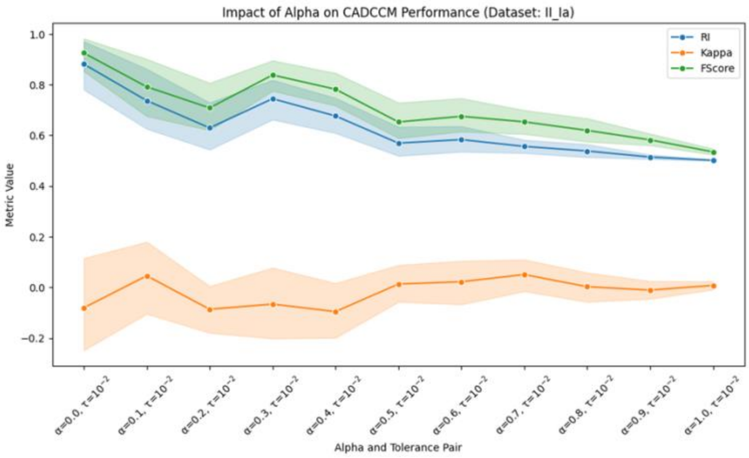
Comparing preprocessing techniques reveals that Gaussian filtering is the most effective method for enhancing EEG clustering performance using CADCCM. It offers the best trade-off between noise reduction, feature preservation, and cluster separability, consistently outperforming other methods in RI, F-score, and Kappa. Despite slightly higher variance, its superior mean scores validate its robustness and reliability for real-world EEG clustering applications.

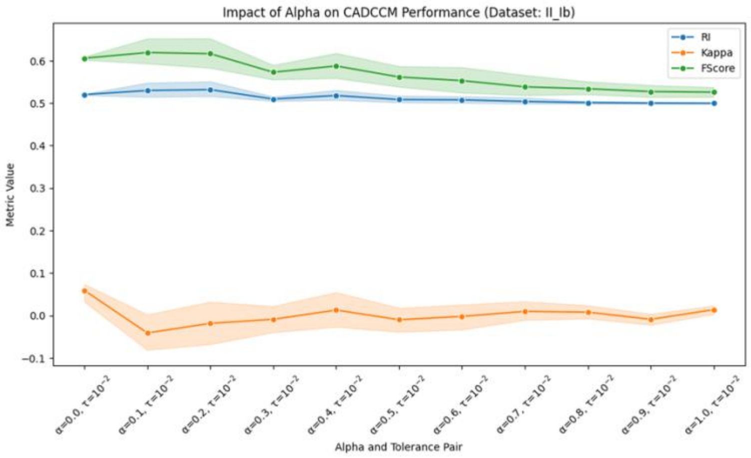
C. PARAMETER SELECTION AND OPTIMIZATION

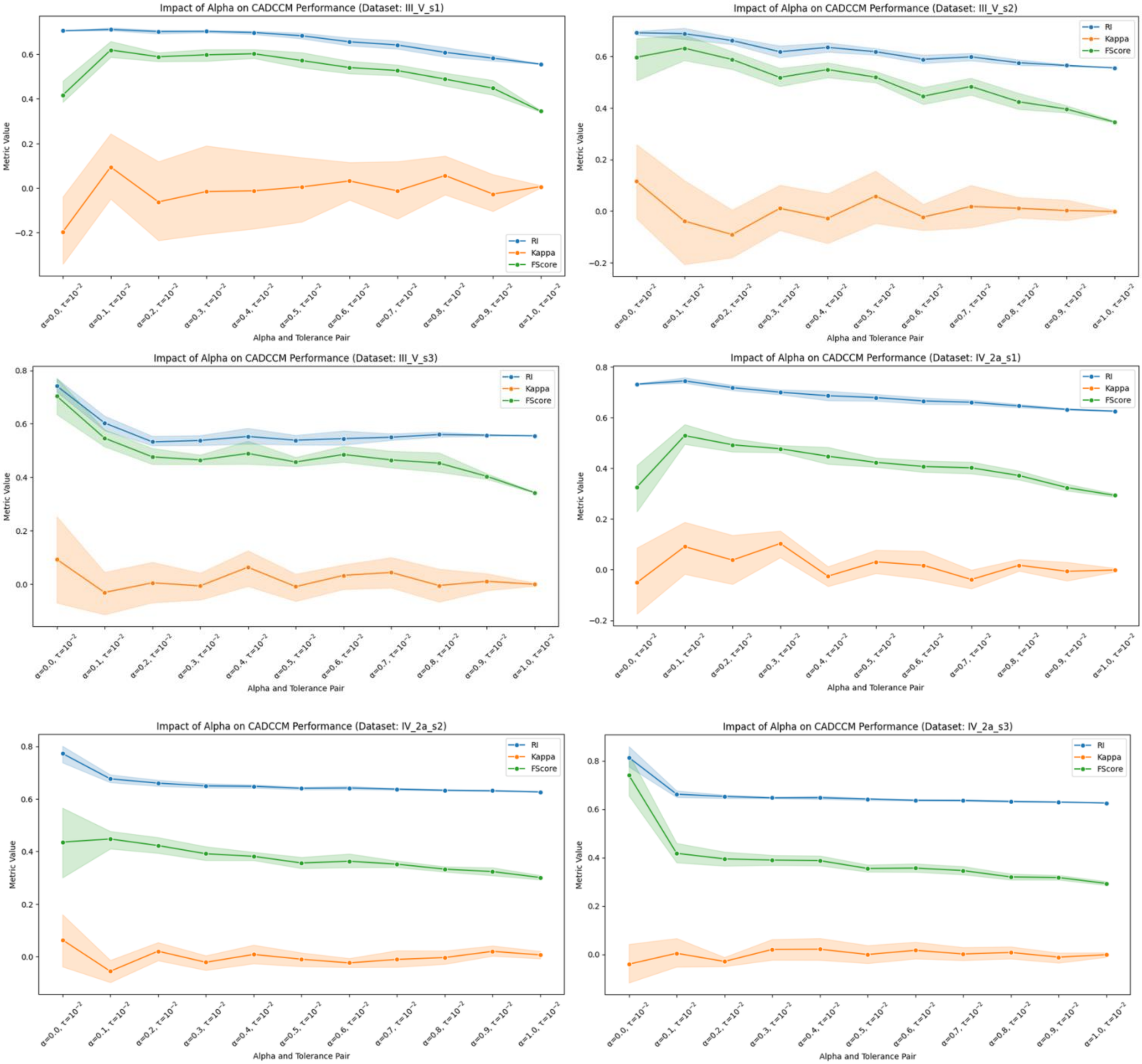
To further enhance CADCCM's performance, an extensive grid search was conducted to optimize alpha, beta, and tolerance values. The impact of these parameters is visualized in Fig. 3, where variations in alpha and tolerance influence clustering accuracy across different datasets. The results indicate that:

- *Higher alpha values,* which emphasize membership, stabilize clustering in datasets with distinct cluster structures.
- *Higher beta values* influence credibility adjustments and improve clustering reliability in datasets with more prevalent noise.
- *Lower tolerance values* yield higher precision at the expense of increased computational cost, while higher values expedite convergence at the cost of clustering accuracy.

Combining Gaussian filtering with an optimized parameter set further boosts CADCCM's clustering accuracy, demonstrating its adaptability across diverse EEG datasets. The findings suggest that proper noise filtering and hyperparameter tuning are crucial in ensuring robust and stable clustering, particularly in biomedical applications where precision is paramount.






Impact of Alpha on CADCCM Performance (Dataset: III_V_s1)
Impact of Alpha on CADCCM Performance (Dataset: III_V_s2)
Impact of Alpha on CADCCM Performance (Dataset: III_V_s3)
Impact of Alpha on CADCCM Performance (Dataset: IV_2a_s1)
Impact of Alpha on CADCCM Performance (Dataset: IV_2a_s2)
Impact of Alpha on CADCCM Performance (Dataset: IV_2a_s3)
RI
Kappa
FScore
Metric Value
Alpha and Tolerance Pair
$\alpha=0.0, \tau=10^{-2}$
$\alpha=0.1, \tau=10^{-2}$
$\alpha=0.2, \tau=10^{-2}$
$\alpha=0.3, \tau=10^{-2}$
$\alpha=0.4, \tau=10^{-2}$
$\alpha=0.5, \tau=10^{-2}$
$\alpha=0.6, \tau=10^{-2}$
$\alpha=0.7, \tau=10^{-2}$
$\alpha=0.8, \tau=10^{-2}$
$\alpha=0.9, \tau=10^{-2}$
$\alpha=1.0, \tau=10^{-2}$
−0.2
0.0
0.2
0.4
0.6
0.8

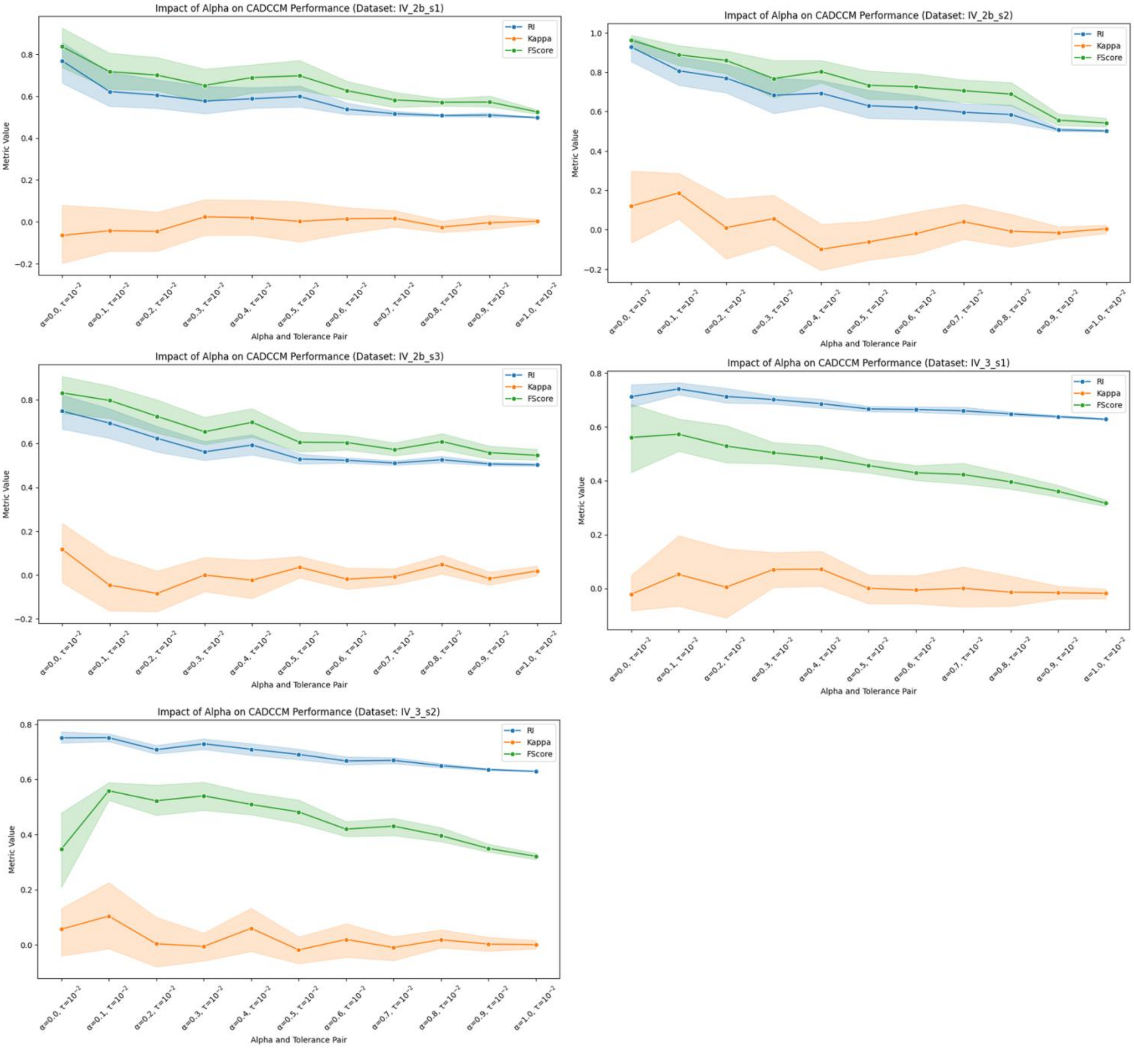


FIGURE 3. **Impact of Alpha, Beta, and Tolerance on CADCCM with Gaussian Filter Across 13 EEG Datasets.**

D. EXECUTION TIME ANALYSIS

The computational efficiency of CADCCM was evaluated across 13 EEG datasets, as shown in Fig. 4. The time complexity of CADCCM is primarily influenced by the iterative updates of membership and credibility scores, making it dependent on the number of clusters (c), data points (n), and iterations (t). Unlike traditional fuzzy clustering methods, CADCCM incorporates credibility adjustments, introducing additional computational overhead. The algorithm was implemented in Python using PyTorch and executed on an iMac with GPU acceleration, significantly improving its efficiency.

As illustrated in Fig. 4, CADCCM requires more computational time than simpler methods, such as K-Means, FCM, K-Medoids, Spectral Clustering, SOM, and Rough Set Clustering (RS), due to its enhanced membership update mechanism and credibility-driven clustering adjustments. However, CADCCM remains computationally feasible for large-scale EEG datasets, providing a favorable trade-off between execution time and clustering accuracy. The grid search optimization for alpha (α), beta (β), and tolerance (τ) further increases the computational load but ensures improved clustering performance. By maintaining both efficiency and accuracy, CADCCM proves to be well-suited for clustering applications involving EEG data.

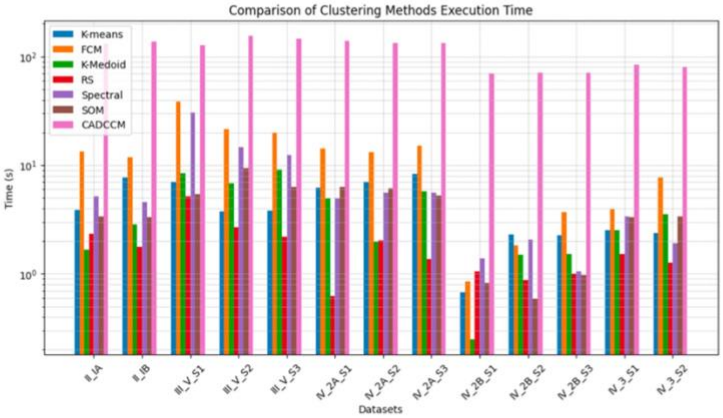


FIGURE 4. Comparison of execution time across clustering methods on EEG datasets.

E. PROPERTIES SATISFIED BY CADCCM

Clustering algorithms are evaluated based on fundamental theoretical properties to ensure their robustness and applicability across diverse datasets. Four key criteria—Scale Invariance, Richness, Consistency, and Order Independence—are essential in assessing a method's effectiveness[40]. Scale Invariance ensures clustering decisions rely on relative rather than absolute distances. Richness guarantees that the algorithm can generate any possible partitioning through parameter tuning. Consistency ensures that changes in intra- and inter-cluster distances do not alter cluster assignments. Finally, Order Independence assesses whether the algorithm's output remains stable regardless of the order in which data points are processed. Table XI analyzes CADCCM's adherence to these criteria.

TABLE XI
ASSESSMENT OF CADCCM BASED ON CLUSTERING CRITERIA

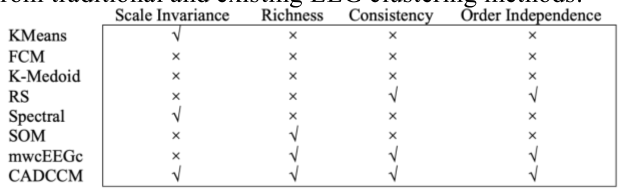
Clustering Criterion	Satisfaction in CADCCM	Justification
Scale Invariance	Yes	Uses credibility-adjusted membership updates rather than absolute distances.
Richness	Yes	Grid search over α, β, and tolerance values enable flexible partitioning.
Consistency	Yes	Clustering assignments remain stable when intra-cluster distances shrink, and inter-cluster distances expand.
Order Independence	Partially	Final clusters stabilize after iterative updates, but initial membership randomness may introduce small variations.

CADCCM fulfills three primary clustering criteria—scale Invariance, Richness, and Consistency—and shows partial independence from data ordering. The minor variation in order dependence is mitigated through iterative updates and parameter tuning, ensuring stable final clustering results. Fig. 5 presents a comparative analysis of clustering properties (scale invariance, richness, consistency, and order independence) across various clustering algorithms, including Rough Set Clustering, Spectral Clustering, K-Means, FCM, K-Medoids, SOM, mwcEEGc, and CADCCM.

- *Scale Invariance:* Methods that do not rely on absolute distance scales satisfy this property. Spectral Clustering, SOM, K-Means, and CADCCM adhere to this criterion, whereas methods like FCM and mwcEEGc depend on predefined similarity thresholds, making them scale-dependent.
- *Richness:* SOM, mwcEEGc, and CADCCM methods are flexible and can generate a variety of cluster partitions based on their parameters. K-Means, Spectral, FCM, and K-Medoid have more rigid constraints on the number of clusters, limiting their richness.
- *Consistency:* Consistency ensures that clusters remain stable when intra-cluster distances shrink and inter-cluster distances expand. Rough Set Clustering, mwcEEGc, and CADCCM exhibit consistency, while centroid-based methods like K-Means, Spectral Clustering, and FCM fail to maintain stability under such conditions.

• *Order Independence:* This property implies that clustering results remain unchanged regardless of the order in which data points are presented. mwcEEGc and CADCCM satisfy this criterion, whereas K-Means, SOM, and other iterative clustering methods produce different results depending on initialization and data order.

Overall, CADCCM demonstrates superior adaptability, satisfying all four key clustering properties, setting it apart from traditional and existing EEG clustering methods.



	Scale Invariance	Richness	Consistency	Order Independence
KMeans	√	×	×	×
FCM	×	×	×	×
K-Medoid	×	×	×	×
RS	×	×	√	√
Spectral	√	×	×	×
SOM	×	√	×	×
mwcEEGc	×	√	√	√
CADCCM	√	√	√	√

FIGURE 5. Comparison of clustering properties across different clustering algorithms. (✓indicates the property is satisfied, while × denotes it is not.)

F. DISCUSSION

1) FINDINGS FROM EXPERIMENTAL EVALUATION

The experimental results demonstrate the effectiveness of CADCCM in EEG clustering compared to existing methods. Key findings include:

- *Clustering Accuracy:* CADCCM consistently outperforms traditional clustering techniques such as K-Means, FCM, K-Medoids, Spectral Clustering, and Rough Set Clustering (RS). It also surpasses the SOM and mwcEEGc methods in several datasets, particularly in cases where the EEG data exhibits high noise levels or overlapping distributions.

- *Impact of Preprocessing:* The choice of preprocessing method significantly affects clustering accuracy. Gaussian filtering consistently led to the highest RI, F-score, and Kappa values, making it the most effective noise reduction technique. It enhances feature separability while preserving crucial EEG signal characteristics.
- *Computational Efficiency:* CADCCM has higher computational time than simpler methods due to credibility adjustments and grid search optimization, but it remains faster than mwcEEGc. GPU acceleration enhances efficiency, making CADCCM feasible for large EEG datasets. The execution time analysis (Fig. 4) demonstrates that CADCCM balances computational cost and clustering accuracy.

2) PERFORMANCE INSIGHTS

Compared to other clustering methods, CADCCM demonstrates robustness in handling noise and non-linear separability. The credibility-adjusted clustering framework refines membership assignment, enhancing cluster stability and accuracy. Integrating effective preprocessing methods, particularly Gaussian filtering, improves performance by mitigating noise-related errors. The results reinforce the importance of parameter tuning and noise reduction in optimizing EEG clustering outcomes.

3) LIMITATIONS & FUTURE DIRECTIONS

Despite its strong performance, CADCCM has areas for improvement. Future research directions include:

- *Deep-Learning-Assisted Feature Extraction:* Integrating deep learning-based feature extraction techniques, such as CNNs, to extract discriminative features before clustering. These methods could improve feature separability and reduce reliance on handcrafted preprocessing techniques. Self-supervised learning approaches may be particularly beneficial for learning robust EEG representations without requiring labeled data.
- *Real-Time Applications:* Investigating the feasibility of CADCCM for real-time EEG clustering is crucial for BCIs and neurological monitoring applications. Future improvements should focus on optimizing computational efficiency by reducing iteration complexity, implementing online or streaming clustering techniques, and leveraging parallel processing with GPUs. These optimizations would enable CADCCM to process continuous EEG streams with minimal latency, making it suitable for real-time decision-making in clinical and assistive technologies.
- *Multi-Modal Data Integration:* Extending CADCCM to integrate multi-modal physiological data (e.g., EEG + ECG) could provide a more comprehensive understanding of brain dynamics. Multi-view clustering techniques can be explored to facilitate the meaningful fusion of heterogeneous signals, thereby improving clustering accuracy.
- *Adaptive Hyperparameter Tuning*: Implementing adaptive optimization strategies to dynamically adjust α, β, and tolerance values based on dataset characteristics could improve clustering adaptability across diverse EEG datasets.

V. CONCLUSION

With the growing volume of unlabeled neurodata, unsupervised clustering remains a complex yet essential task. This study introduced the Credibility-Adjusted Data-Conscious Clustering Method (CADCCM), a novel approach for unsupervised EEG clustering that integrates credibility adjustments with fuzzy membership updates. CADCCM achieves superior clustering performance across multiple EEG datasets by incorporating Gaussian-based noise filtering and leveraging GPU acceleration. Experimental results demonstrate that CADCCM consistently outperforms traditional clustering methods, including K-Means, FCM, K-Medoids, Spectral, Rough Set clustering, SOM, and mwEEGc, particularly in datasets with overlapping distributions and high artifact contamination. Furthermore, CADCCM demonstrates superior performance by satisfying all key clustering properties, including scale invariance, richness, and consistency, reinforcing its theoretical soundness.

The proposed method effectively balances clustering accuracy and computational efficiency, with credibility-based adjustments enhancing cluster separability while maintaining a feasible runtime for large-scale EEG datasets. Although CADCCM requires additional computational resources compared to simpler methods, GPU acceleration significantly reduces execution time, making it scalable for real-world EEG applications.

Future work will focus on adaptive hyperparameter tuning to optimize performance across diverse EEG datasets. Additionally, extending CADCCM to multi-modal physiological data and exploring its integration with deep-learning-assisted feature extraction can further enhance its applicability in brain-computer interfaces (BCIs), cognitive state monitoring, and neurological diagnostics. Furthermore, investigating real-time applications of CADCCM could provide valuable insights into its effectiveness for continuous EEG signal analysis in medical and cognitive research.

REFERENCES

[1] C. Dai *et al.*, "Brain EEG Time-Series Clustering Using Maximum-Weight Clique," *IEEE Trans Cybern*, vol. 52, no. 1, pp. 357–371, Jan. 2022, doi: 10.1109/TCYB.2020.2974776.

[2] A. S.A *et al.*, "Analysis of EEG microstates as biomarkers in neuropsychological processes – Review," *Comput Biol Med*, vol. 173, p. 108266, May 2024, doi: 10.1016/J.COMPBIOMED.2024.108266.

[3] T. Y. Wen and S. A. Mohd Aris, "Hybrid Approach of EEG Stress Level Classification Using K-Means Clustering and Support Vector Machine," *IEEE Access*, vol. 10, pp. 18370–18379, 2022, doi: 10.1109/ACCESS.2022.3148380.

[4] Y. Zhang *et al.*, "Improving EEG Decoding via Clustering-Based Multitask Feature Learning," *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 8, pp. 3587–3597, Aug. 2022, doi: 10.1109/TNNLS.2021.3053576.

[5] Z. A. A. Alyasseri *et al.*, "EEG-Based Person Identification Using Multi-Verse Optimizer as Unsupervised Clustering Techniques," *Evolutionary Data Clustering: Algorithms and Applications*, pp. 89–110, 2021, doi: 10.1007/978-981-33-4191-3_4.

[6] R. Srinath, R. Gayathri, C. Shalini, and P. Maragathavalli, "Epilepsy Disease Detection Using the Proposed CNN-FCM Approach," in *International Conference on Innovations in Data Analytics*, Springer, Singapore, 2024, pp. 371–380. doi: 10.1007/978-981-97-4928-7_29.

[7] P. Pratyasha and S. Gupta, "Obstructive Sleep Apnea Detection from EEG Data: A Hybrid Approach of One-Dimensional Convolutional Neural Network and Enhanced Fuzzy C-Means Clustering Algorithm," *Sleep Vigil*, vol. 8, no. 2, pp. 231–243, Dec. 2024, doi: 10.1007/S41782-024-00282-7/METRICS.

[8] S. Kumar Reddy C and S. M, "A 1-D CNN-FCM model for the classification of epileptic seizure disorders," *Neural Comput Appl*, vol. 35, no. 24, pp. 17871--17881, Aug. 2023.

[9] P. K. Upadhyay and C. Nagpal, "PCA-Aided FCM Identifies Stressful Events of Sleep EEG Under Hot Environment," *IETE J Res*, vol. 68, no. 5, pp. 3862–3875, Sep. 2022, doi: 10.1080/03772063.2020.1782273.

[10] C. Zhao *et al.*, "Leveraging Large Language Models and Fuzzy Clustering for EEG Report Analysis," *Proceedings of IEEE Sensors*, 2024, doi: 10.1109/SENSORS60989.2024.10784894.

[11] M. A. Li, R. T. Wang, and L. N. Wei, "Fuzzy support vector machine with joint optimization of genetic algorithm and fuzzy c-means," *Technology and Health Care*, vol. 29, no. 5, pp. 921–937, Jan. 2021, doi: 10.3233/THC-202619.

[12] E. Thomas and S. N. Kumar, "Fuzzy C Means Clustering Coupled with Firefly Optimization Algorithm for the Segmentation of Neurodisorder Magnetic Resonance Images," *Procedia Comput Sci*, vol. 235, pp. 1577–1589, Jan. 2024, doi: 10.1016/J.PROCS.2024.04.149.

[13] G. Narula, M. Haeberlin, J. Balsiger, C. Strässle, L. L. Imbach, and E. Keller, "Detection of EEG burst-suppression in neurocritical care patients using an unsupervised machine learning algorithm," *Clinical Neurophysiology*, vol. 132, no. 10, pp. 2485–2492, Oct. 2021, doi: 10.1016/J.CLINPH.2021.07.018.

[14] N. Sukhorukova, J. A. M. E. S. Willard-Turton, G. Garwoli, C. Morgan, and A. Rokey, "Spectral Clustering and Long Timeseries Classification," *The ANZIAM Journal*, vol. 66, no. 2, pp. 121–131, 2024, doi: 10.1017/S1446181124000105.

[15] Y. Zhang, Y. Jiang, L. Qi, M. Z. A. Bhuiyan, and P. Qian, "Epilepsy Diagnosis Using Multi-view & Multi-medoid Entropy-based Clustering with Privacy Protection," *ACM Trans Internet Technol*, vol. 21, no. 2, May 2021, doi: 10.1145/3404893.

[16] G. H. Martono and N. Sulistianingsih, "Optimizing Autism Spectrum Disorder Identification with Dimensionality Reduction Technique and K-Medoid," *JURNAL INFOTEL*, vol. 16, no. 4, pp. 837-854–837–854, Dec. 2024, doi: 10.20895/INFOTEL.V16I4.1142.

[17] Q. Ren, X. Sun, X. Fu, S. Kumar Arjaria, G. Chaubey, and N. Shukla, "Hjorth Parameter based Seizure Diagnosis using Cluster Analysis," *J Phys Conf Ser*, vol. 1998, no. 1, p. 012020, Aug. 2021, doi: 10.1088/1742-6596/1998/1/012020.

[18] I. Exarchos *et al.*, "Supervised and unsupervised machine learning for automated scoring of sleep-wake and cataplexy in a mouse model of narcolepsy," *Sleep*, vol. 43, no. 5, May 2020, doi: 10.1093/SLEEP/ZSZ272.

[19] Y. Lee and M. Alamaniotis, "Unsupervised EEG Cybersickness Prediction with Deep Embedded Self Organizing Map," *Proceedings - IEEE 20th International Conference on Bioinformatics and Bioengineering, BIBE 2020*, pp. 538–542, Oct. 2020, doi: 10.1109/BIBE50027.2020.00093.

[20] X. Wan, Z. Fang, M. Wu, and Y. Du, "Automatic detection of HFOs based on singular value decomposition and improved fuzzy c-means clustering for localization of seizure onset zones," *Neurocomputing*, vol. 400, pp. 1–10, Aug. 2020, doi: 10.1016/J.NEUCOM.2020.03.010.

[21] A. Lekova and I. Chavdarov, "A Fuzzy Shell for Developing an Interpretable BCI Based on the Spatiotemporal Dynamics of the Evoked Oscillations," *Comput Intell Neurosci*, vol. 2021, no. 1, p. 6685672, Jan. 2021, doi: 10.1155/2021/6685672.

[22] T. Gao *et al.*, "Adaptive density peaks clustering: Towards exploratory EEG analysis," *Knowl Based Syst*, vol. 240, p. 108123, Mar. 2022, doi: 10.1016/J.KNOSYS.2022.108123.

[23] Y. Zhang *et al.*, "Unsupervised Time-Aware Sampling Network with Deep Reinforcement Learning for EEG-Based Emotion Recognition," *IEEE Trans Affect Comput*, Dec. 2022, doi: 10.1109/TAFFC.2023.3319397.

[24] S. H. Hsu, Y. Lin, J. Onton, T. P. Jung, and S. Makeig, "Unsupervised learning of brain state dynamics during emotion imagination using high-density EEG," *Neuroimage*, vol. 249, p. 118873, Apr. 2022, doi: 10.1016/J.NEUROIMAGE.2022.118873.

[25] İ. Yıldız, R. Garner, M. Lai, and D. Duncan, "Unsupervised seizure identification on EEG," *Comput Methods Programs Biomed*, vol. 215, p. 106604, Mar. 2022, doi: 10.1016/J.CMPB.2021.106604.

[26] S. Pan, T. Shen, Y. Lian, and L. Shi, "A Task-Related EEG Microstate Clustering Algorithm Based on Spatial Patterns, Riemannian Distance, and a Deep Autoencoder," *Brain Sciences 2025, Vol. 15, Page 27*, vol. 15, no. 1, p. 27, Dec. 2024, doi: 10.3390/BRAINSCI15010027.

[27] T. Kawaguchi, K. Ono, and H. Hikawa, "Electroencephalogram-Based Facial Gesture Recognition Using Self-Organizing Map," *Sensors 2024, Vol. 24, Page 2741*, vol. 24, no. 9, p. 2741, Apr. 2024, doi: 10.3390/S24092741.

[28] X. Zhu, C. Liu, L. Zhao, and S. Wang, "EEG Emotion Recognition Network Based on Attention and Spatiotemporal Convolution," *Sensors 2024, Vol. 24, Page 3464*, vol. 24, no. 11, p. 3464, May 2024, doi: 10.3390/S24113464.

[29] Y. Zhao *et al.*, "Latent Prototype-Based Clustering: A Novel Exploratory Electroencephalography Analysis Approach," *Sensors 2024, Vol. 24, Page 4920*, vol. 24, no. 15, p. 4920, Jul. 2024, doi: 10.3390/S24154920.

[30] Y. Murali Krishna and P. Vinay Kumar, "Efficient automated method to extract EOG artifact by combining Circular SSA with wavelet and unsupervised clustering from single channel EEG," *Biomed Signal Process Control*, vol. 87, p. 105455, Jan. 2024, doi: 10.1016/J.BSPC.2023.105455.

[31] X. Qu and T. J. Hickey, "EEG4Home: A Human-In-The-Loop Machine Learning Model for EEG-Based BCI," in *International Conference on Human-Computer Interaction*, Springer, Cham, 2022, pp. 162–172. doi: 10.1007/978-3-031-05457-0_14.

[32] J. Cao *et al.*, "Unsupervised Eye Blink Artifact Detection From EEG With Gaussian Mixture Model," *IEEE J Biomed Health Inform*, vol. 25, no. 8, pp. 2895–2905, Aug. 2021, doi: 10.1109/JBHI.2021.3057891.

[33] T. J. Loftus *et al.*, "Unsupervised EEG Artifact Detection and Correction," *Front Digit Health*, vol. 2, p. 608920, Jan. 2021, doi: 10.3389/FDGTH.2020.608920.

[34] Y. Du, G. Li, M. Wu, and F. Chen, "Unsupervised Multivariate Feature-Based Adaptive Clustering Analysis of Epileptic EEG Signals," *Brain Sciences 2024, Vol. 14, Page 342*, vol. 14, no. 4, p. 342, Mar. 2024, doi: 10.3390/BRAINSCI14040342.

[35] R. Mahini *et al.*, “Brain Evoked Response Qualification Using Multi-Set Consensus Clustering: Toward Single-Trial EEG Analysis,” *Brain Topogr*, vol. 37, no. 6, Nov. 2024, doi: 10.1007/S10548-024-01074-Y.

[36] M. M. Shaker, “EEG Waves Classifier using Wavelet Transform and Fourier Transform,” *World Academy of Science, Engineering and Technology, International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering*, vol. 1, no. 3, pp. 169.-174, 2007.

[37] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, A. K. Abasi, and S. N. Makhadmeh, “EEG Signals Denoising Using Optimal Wavelet Transform Hybridized with Efficient Metaheuristic Methods,” *IEEE Access*, vol. 8, pp. 10584–10605, 2020, doi: 10.1109/ACCESS.2019.2962658.

[38] M. Grobbelaar *et al.*, “A Survey on Denoising Techniques of Electroencephalogram Signals Using Wavelet Transform,” *Signals 2022, Vol. 3, Pages 577-586*, vol. 3, no. 3, pp. 577–586, Aug. 2022, doi: 10.3390/SIGNALS3030035.

[39] A. Harishvijey and J. Benadict Raja, “Automated technique for EEG signal processing to detect seizure with optimized Variable Gaussian Filter and Fuzzy RBFELM classifier,” *Biomed Signal Process Control*, vol. 74, p. 103450, Apr. 2022, doi: 10.1016/J.BSPC.2021.103450.

[40] J. Kleinberg, “An Impossibility Theorem for Clustering,” *Adv Neural Inf Process Syst*, vol. 15, 2002.

F. DIVAN Fatemeh Divan is a Ph.D. candidate at the University of Malaya (UM). She earned her MSc in Computer Science in 2015 and her BSc in Computer Science in 2013, both from the University of Malaya. Additionally, she holds a BSc in Pure Mathematics, obtained in 2008 from K.N. Toosi University of Technology in Iran. Her master’s dissertation focused on developing a hybrid evolutionary algorithm for solving constrained engineering problems. Her research interests include artificial intelligence, machine learning, robotics, and mathematics.

T. Y. Wah Ying Wah Teh received his Ph.D. in data warehousing from the University of Malaya, Kuala Lumpur, Malaysia, in 2004. He is currently a professor in the information science department of the Faculty of Computer Science and Information Technology at the University of Malaya. His research interests include data mining and text mining.

K. S. Lim Professor Dr. Kheng Seang Lim, a University of Malaya (Malaysia) graduate [1999] with MRCP [2004] and FRCP, is a Professor of Neurology in the Faculty of Medicine, University of Malaya (UM) and Consultant Neurologist specializing in epilepsy in University of Malaya Medical Centre and University Malaya Specialist Centre, Malaysia. He has been trained in the UM for his neurology subspecialty training [2008], followed by fellowship training in Melbourne [2011] and Cleveland [2017] for epilepsy. He completed his junior doctor training in Hospital Alor Setar, Kedah, before joining the University Malaya Medical Centre in 2005. He is currently the Associate Editor for Neurology Asia, an Editorial Board member of the Journal of Xiangya Medicine, and a committee member of the recent 7th Asian Oceanian Congress on Clinical Neurophysiology (AOCCN) held in February 2021. He has published ≥ 120 original papers with an H-index of 16, on the psychosocial aspects of epilepsy, genetics, pharmacogenomics, pharmacokinetics of antiepileptic drugs, quantitative EEG, and quantitative MRI research. He is actively involved in research as the principal investigator and co-investigator in numerous local and international research projects. He has established strong collaborations with Taiwan, New Zealand, Singapore, and the United Kingdom. He was recently awarded the Impact-Oriented Interdisciplinary Research Grant (IIRG) Programme for epilepsy research. His areas of expertise and research interest include Neurology, Epilepsy, Epilepsy Surgery, Electroencephalography, Psychosocial Research in Epilepsy, Genetics in Familial Epilepsy, Population Genetics, and Molecular Pharmacology. He is an active member of the Malaysian Epilepsy Council and the past President of the Malaysian Epilepsy Society and Malaysian Society of Neurosciences. He was chair of the Research Commission in the International Bureau of Epilepsy and the Research Task Force in the Commission of Asian and Oceanian Affairs, International League Against Epilepsy.

A. S. Shirkhorshidi Ali Seyed Shirkhorshidi earned his Ph.D. in Machine Learning from the University of Malaya in 2020, after completing a master’s in software engineering from Staffordshire University in 2013 and a bachelor’s in computer science from Payam Noor University in 2010. With over a decade of experience in academia and industry, he has led numerous projects. Currently, he serves as the Director of Data Science at Entefy, where he drives innovation by applying advanced data science and machine learning techniques.