

Tensor decomposition-based feature extraction for EEG signals in epileptic seizure detection

Raul Victor de O. Paiva, Alla Manina,

Tarcisio F. Maciel, *Senior Member, IEEE*, and Martin Haardt, *Fellow, IEEE*

Abstract—Epilepsy monitoring based on electroencephalogram (EEG) signals is important for detecting pre-seizure states, but multidimensional EEG representations can lead to high-dimensional feature spaces and increased computational cost. This work proposes two canonical polyadic decomposition (CPD)-based approaches for extracting compact features from spectral-temporal-spatial EEG tensors obtained with a continuous wavelet transform (CWT). The proposed approaches use the estimated low-rank factor matrices to preserve physically interpretable information while reducing the number of features used to train machine learning (ML) classifiers. Experimental results show robust classification performance for inter- and pre-ictal classification, with performance metrics of up to approximately 99%, while substantially reducing the feature-space dimensionality under patient-specific variability.

Keywords—EEG, CPD, feature extraction, machine learning.

I. INTRODUCTION

Epilepsy is a chronic, life-threatening neurological disorder. Anticipating seizures via electroencephalogram (EEG) analysis allows patients to take precautionary measures against accidents [1]. As shown in Fig. 1, the epilepsy timeline comprises four phases according to [1]: ictal (red, the seizure), pre-ictal (orange, immediately before the seizure), post-ictal (blue, after the seizure), and inter-ictal (green, between seizures). The seizure prediction horizon (SPH) (dashed red) acts as a safety margin to prevent leakage of epileptic waves during seizures and to create a safe time margin for the detection of imminent epileptic seizures.

In [1], inter-ictal intervals are selected from periods after a 30-minutes post-ictal period to ensure they represent normal brain activity. Meanwhile, the duration of pre-ictal windows can be set as 5, 30, 120 or 240 minutes prior to an ictal phase, enabling the detection of changes within these windows.

Recent studies have shown the use of tensor decompositions (TDs) for EEG signals feature extraction in classification tasks, preserving interpretability [2]–[4]. In [2], a canonical polyadic decomposition (CPD) and a block-term decomposition (BTD) are applied to low-rank EEG tensors from the Cuban neurological restoration center (CIREN) epilepsy unit to detect non-convulsive epileptic seizures. The three-dimensional (spectral-temporal-spatial) tensors are built using both a continuous wavelet transform (CWT) and Hilbert-Huang (HH) techniques.

R. V. de O. Paiva and T. F. Maciel, Post-Graduation Program in Teleinformatics Engineering, UFC, Fortaleza, Brazil, e-mail: raul.paiva@alu.ufc.br, maciel@ufc.br; Alla Manina and Martin Haardt, Communications Research Laboratory, TU Ilmenau, Ilmenau, Germany, e-mail: {alla.manina, martin.haardt}@tu-ilmenau.de. This work was partially supported by INCT Signals (grant INCT-25255-82587.32.41/64-FUNCAP, 406517/2022-3-CNPq), and CAPES Finance code 001. T. F. Maciel and R. V. de O. Paiva were partially supported by CAPES-PROBRAL (grant 88881.894956/2023-01), FUNCAP (no. BMD-0008-00514.01.23/24) and CNPq (grants 312471/2021-1 and 305165/2026-7). A. Manina was supported by the Carl-Zeiss-Stiftung under the “PollenNet” project (no. P2022-08-006).

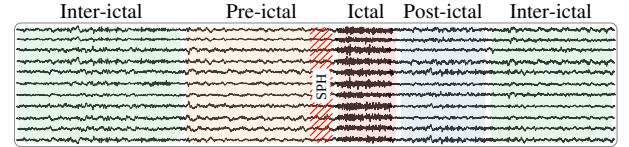


Fig. 1. Illustration of the epileptic seizure intervals.

Among several evaluated classifiers fed with vectorized factor matrices, the combination of CPD features from HH tensors and a support vector machine (SVM)-radial basis function (RBF) model yielded the best results.

In [3], the simultaneous multilinear low-rank approximation of tensors approach and its non-negative variant use Tucker decomposition on 4th-order EEG tensors (spectral-temporal-spatial-sample) for epileptic spike detection. Here, the sample domain defines a feature space governed solely by the multilinear rank, allowing new test samples to be projected onto the training space for feature extraction. Evaluated on the Children’s hospital Boston (CHB)-Massachusetts Institute of Technology (MIT) dataset across several machine learning (ML) models, a SVM achieved optimal performance, yielding an average sensitivity of $\approx 80\%$, specificity $> 90\%$, and high area under the receiver operating characteristic (ROC) curve (AUC) values. However, because this framework relies on simultaneous decomposition, it introduces high storage demands that limit its scalability to the large-scale EEG datasets.

For Alzheimer’s dementia detection, [4] applied CPD to 3rd-order EEG tensors spanning frequency, spatial, and recording segment domains, using the power spectral density (PSD) of each recording segment to classify cognitive impairment stages. Following the projection methodology of [3], this approach differs primarily by utilizing a 3rd instead of a 4th-order tensor structure. Evaluated on the Temple University Hospital (TUH)-EEG corpus using various ML models, the resulting low-rank factors yielded accurate stage classification, achieving AUC values of $\approx 92\%$ with Gaussian naive Bayes (GNB) and SVM classifiers. Like [3], this approach struggles with simultaneous large-scale dataset decomposition.

Recent studies leverage deep learning (DL) to automatically extract features from EEG signals for epileptic seizure detection [5]–[7]. Specifically, [5] proposes a hybrid convolutional neural network (CNN)-Transformer model, while [6] employs a synchronization-based graph spatio-temporal attention network. Meanwhile, [7] introduces a 3D-squeeze-and-excitation residual network (SERESNet) framework designed for both patient-specific and multi-patient generalization.

The main objective of this work is to propose a CPD-based methodology for EEG feature extraction and selection. The proposed methodology exploits low-rank representations of spectral-temporal-spatial EEG tensors to obtain compact and physically interpretable features for epileptic seizure detection.

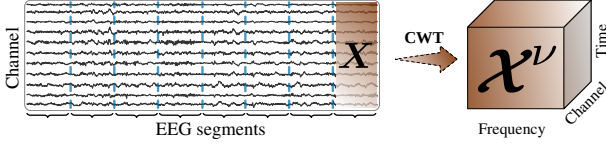


Fig. 2. EEG segment in its matrix and tensor forms.

In this paper, CPD via SEmi-Algebraic framework for approximate CP decomposition via Simultaneous matrix diagonalization (SECSI) [8] is used for this purpose. The estimated factor matrices, associated with the spectral, temporal, and spatial dimensions, are used to construct feature vectors for training a linear SVM to classify inter- and pre-ictal segments. The remainder of this paper reviews CPD-based EEG feature extraction in Section II, presents the proposed approaches in Section III, reports the dataset, metrics, methodology, and results in Section IV, and Section V concludes our findings.

II. CPD FOR EEG SIGNAL FEATURE EXTRACTION

EEG signals record brain electrical activity and are widely used to analyze neural dynamics under different physiological and pathological conditions. These signals exhibit an inherent multidimensional structure, commonly described across time, frequency, and spatial domains, and can therefore be naturally represented as tensors. In the frequency domain, EEG activity is typically divided into five main bands: delta δ (0.3–3 Hz), theta θ (4–7 Hz), alpha α (8–15 Hz), beta β (16–30 Hz), and gamma γ (31–120 Hz), which are associated with different brain states and activities [9].

Consider an EEG signal originally obtained from a time-series segment and represented as a time-channel matrix $\mathbf{X} \in \mathbb{R}^{T \times Ch}$. For each channel, the signal is transformed into a time-frequency representation by applying the CWT over F frequency samples within a frequency band ν , resulting in the 3rd order tensors

$$\mathbf{X}^\nu \in \mathbb{C}^{F \times T \times Ch}, \quad \nu \in \{\delta, \theta, \alpha, \beta, \gamma\}. \quad (1)$$

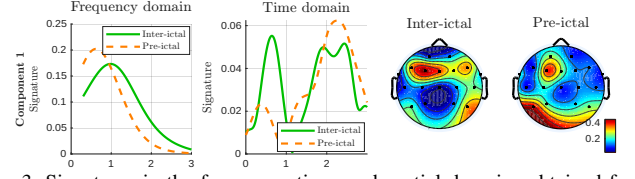
Fig. 2 shows the representation of an EEG segment as both a time-channel matrix and a multidimensional tensor.

Since the CWT-based tensor representation captures information across the frequency, time, and spatial domains, the vectorized form of $|\mathbf{X}^\nu|$, $\mathbf{x}^\nu \in \mathbb{R}^{F \cdot T \cdot Ch \times 1}$, can be used as a feature vector for classification. This representation yields $F \cdot T \cdot Ch$ features, leading to a high-dimensional feature space that scales with the tensor dimensions. Therefore, data compression methods that preserve essential information from brain signals are desirable for feature extraction and subsequent training of ML models. In this context, TD approaches can be used to extract low-rank features from multidimensional EEG data [4].

The CPD provides a structured low-rank approximation of multidimensional data [10], and we apply it here for feature extraction from EEG tensors. For a 3rd order tensor $\mathbf{X}^\nu$, $\nu \in \{\delta, \theta, \alpha, \beta, \gamma\}$, the CPD is defined as a sum of R rank-one tensor components, i.e.,

$$\mathbf{X}^\nu \approx \sum_{r=1}^R \mathbf{a}_r^\nu \circ \mathbf{b}_r^\nu \circ \mathbf{c}_r^\nu \in \mathbb{C}^{F \times T \times Ch}, \quad (2)$$

where $\mathbf{a}_r^\nu \in \mathbb{C}^{F \times 1}$, $\mathbf{b}_r^\nu \in \mathbb{C}^{T \times 1}$, and $\mathbf{c}_r^\nu \in \mathbb{C}^{Ch \times 1}$ denote the r -th column of the factor matrices $\mathbf{A}^\nu \in \mathbb{C}^{F \times R}$, $\mathbf{B}^\nu \in \mathbb{C}^{T \times R}$, and $\mathbf{C}^\nu \in \mathbb{C}^{Ch \times R}$, i.e., $\mathbf{A}^\nu = [\mathbf{a}_1^\nu \mathbf{a}_2^\nu \dots \mathbf{a}_R^\nu]$, $\mathbf{B}^\nu = [\mathbf{b}_1^\nu \mathbf{b}_2^\nu \dots \mathbf{b}_R^\nu]$, and $\mathbf{C}^\nu = [\mathbf{c}_1^\nu \mathbf{c}_2^\nu \dots \mathbf{c}_R^\nu]$, respectively. Eq. (2) can also be


 Fig. 3. Signatures in the frequency, time, and spatial domains obtained from a rank-one decomposition of one EEG segment from a subject in the theta band θ , for both inter-ictal (green) and pre-ictal (orange) phases.

reformulated using the n -mode tensor-matrix product ($\times_n$), $n = 1, 2, 3$, and the 3rd order super-diagonal tensor $\mathbf{I} \in \mathbb{R}^{R \times R \times R}$, depicted in the Step 1) of Fig. 4, as

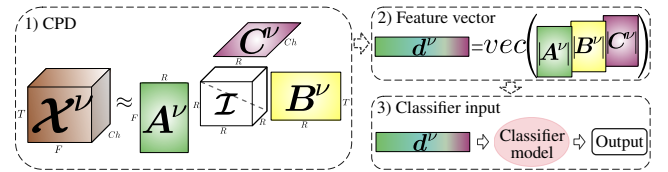
$$\mathbf{X}^\nu \approx \mathbf{I}_{3,R} \times_1 \mathbf{A}^\nu \times_2 \mathbf{B}^\nu \times_3 \mathbf{C}^\nu. \quad (3)$$

According to [4], each component may correspond to a different brain source, which can be associated with noise or epileptic waveforms. In this case, the frequency-domain component signatures describe the relative distribution of wavelet coefficient magnitude across the analyzed frequency bins of a corresponding ν band, $\nu \in \{\delta, \theta, \alpha, \beta, \gamma\}$, whereas the temporal and spatial signatures describe when and where the corresponding activity is expressed.

Suppose that epileptic seizure detection is addressed as a classification task using EEG segment tensors constructed from the CWT in the theta band θ as samples. Considering a sample tensor, Fig. 3 shows the frequency, time and spatial signatures of each component from a rank-one decomposition, for both inter- and pre-ictal phases.

In this example, differences between the inter- and pre-ictal phases can be observed. To evaluate the separability strength of these features, the Fisher score (FS) analysis is performed in Section IV-D considering multiple samples. This analysis investigates whether the intrinsic low-rank structure of the EEG tensors captured by the CPD provides discriminative features for classification. Therefore, the factor matrices $\mathbf{A}^\nu$, $\mathbf{B}^\nu$ and $\mathbf{C}^\nu$ can be explored as inputs to a classifier instead of using all the elements of $\mathbf{X}^\nu$.

By staking and vectorizing the absolute values of the estimated factor matrices, the feature vector $\mathbf{d}^\nu \in \mathbb{R}^{(F+T+Ch) \times 1}$ is constructed. The resulting number of the training features is $R(F+T+Ch)$, which is lower than $(F \cdot T \cdot Ch)$, obtained from the original tensor $\mathbf{X}^\nu$, depending on the selected approximation rank R . Fig. 4 depicts the basic feature extraction process from a 3rd-order EEG tensor using CPD.


 Fig. 4. Basic feature extraction process from a 3rd order EEG tensor using CPD: 1) the estimated factor matrices are used to construct 2) the feature vector $\mathbf{d}^\nu$, which is then used to 3) feed a classifier.

III. EEG-BASED EPILEPTIC SEIZURE DETECTION USING CPD-DERIVED FEATURES

Based on the CWT and CPD results, this section introduces two proposed methods for feature extraction and selection for the inter- and pre-ictal classification task.

The samples correspond to segments extracted from the EEG time series and used to feed the classifier. For each segment, the CWT is computed in parallel for each frequency band ν , yielding 5 tensors $\mathbf{X}^\nu$, with $\nu \in \delta, \theta, \alpha, \beta$, and γ . For a manually selected low rank R , the CPD is then applied to each $\mathbf{X}^\nu$ to estimate the factor matrices $\mathbf{A}^\nu$, $\mathbf{B}^\nu$, and $\mathbf{C}^\nu$, which are then used to construct the feature vectors.

The FS is used to evaluate the separability of the extracted features [11] via the two proposed approaches. Given a set of labeled samples $\{\mathbf{E}_j, y_j\}_{j=1}^n$, with $y_j \in \{1, \dots, k\}$, let n_i denote the number of samples in class i . Let μ_{ip} and σ_{ip}^2 denote the mean and the variance of the p -th feature in class i , respectively, and let μ_p denote the global mean of feature p . The FS is then defined as

$$\mathcal{F}_p = \left(\sum_{i=1}^k n_i (\mu_{ip} - \mu_p)^2 \right) / \left(\sum_{i=1}^k n_i \sigma_{ip}^2 \right). \quad (4)$$

FS ranks features objectively by quantifying the inter-to-intra-class variance ratio, where higher values denote stronger discriminative power [11]. We apply FS across both proposed approaches.

A. Approach 1 (A1): Stacking frequency- and spatial-domain factor matrices across the EEG bands

As will be shown in Section IV-D, the features associated with the temporal factor matrix $\mathbf{B}^\nu$ present low FSs, indicating overlapping distributions between the inter- and pre-ictal phases and limited relevance for the classification task addressed in this work. Therefore, A1 follows the basic CPD-based feature extraction scheme described in Section II, but discards $\mathbf{B}^\nu$, for all EEG bands ν .

For a given segment sample, the tensors $\mathbf{X}^\nu$, $\nu \in \delta, \theta, \alpha, \beta$, and γ , are first decomposed using CPD. The estimated factor matrices $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ are then stacked and vectorized to construct the feature vector, yielding $5R(F+Ch)$ training features. Using all samples, the FS of each feature is computed offline, and the features are reordered from highest to lowest according to their FSs.

Finally, the reordered feature vectors are used as inputs to the classifier to distinguish between inter- and pre-ictal phases. Fig. 5 summarizes the A1.

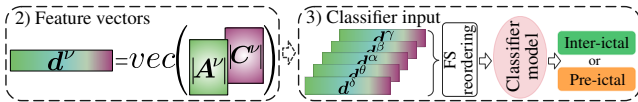


Fig. 5. Approach A1 decomposes EEG tensors (ν) using CPD. 2) Stacking factor matrices $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ yields 3) feature vectors $\mathbf{d}^\nu$, ranked by FS for classification.

B. Approach 2 (A2): Spectral-spatial signature constructed with $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ across EEG bands

The main difference in A2 lies in the construction of the feature vectors. In this approach, the factor matrices $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ are multiplied to obtain the spectral-spatial signature $\mathbf{D}^\nu$, which is then vectorized to form $\mathbf{d}^\nu$ as follows

$$\mathbf{D}^\nu = \mathbf{A}^\nu (\mathbf{C}^\nu)^H \in \mathbb{C}^{F \times Ch}, \quad (5)$$

$$\mathbf{d}^\nu = \text{vec}(|\mathbf{D}^\nu|) \in \mathbb{R}^{FCh \times 1}. \quad (6)$$

The remaining steps, including the reordering of the features according to their FSs, follow the same procedure as in A1. In

this case, the dimensionality of $\mathbf{d}^\nu$ is independent of the rank R , yielding $5(F \cdot Ch)$ features for training the ML models. If $\mathbf{D}^\nu$ is interpreted as an image, the rank R controls the level of spectral-spatial detail captured by this representation. Fig. 6 depicts the step 2) of the proposed A2, followed by the reordering of the extracted feature vectors before being used as classifier inputs.

$$\text{2) Feature vectors} \quad \mathbf{d}^\nu = \left| \text{vec}(\mathbf{A}^\nu \cdot \mathbf{C}^{\nu H}) \right|$$

Fig. 6. Approach A2 decomposes 3rd order EEG tensors (ν) using CPD. 2) The product of $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ forms spectral-spatial vectors $\mathbf{d}^\nu$, ranked by FS for classification.

IV. PERFORMANCE EVALUATION

This section evaluates the proposed approaches.

A. Dataset description

Classification experiments were conducted using the first six subjects (denoted as subjects 1, 2, ..., 6) from the CHB-MIT database, which contains scalp EEG recordings acquired at 256 Hz using the 10-20 system [12]. Following [1], the pre-ictal state is defined as a 30-minute window (600 segments of 3 seconds). To account for event rarity and maintain a balanced classification setting, an equal number of inter-ictal segments was selected for each subject, yielding 1200 samples per subject (600 inter-ictal and 600 pre-ictal) for testing, as summarized in Table I.

Using patient-specific leave-one-seizure-out cross validation (CV) (LOSOCV), each seizure was tested individually while training on the rest, preventing data leakage across events. A broader evaluation with all eligible database subjects is left for future work.

B. Evaluation Metrics

Classification performance is evaluated using recall, precision, and AUC. Let TP , TN , FP , and FN denote the true positives, true negatives, false positives, and false negatives, respectively. Recall measures the proportion of correctly identified pre-ictal segments among all actual pre-ictal segments and is defined as $Rec = [TP/(TP+FN)] \cdot 100\%$. A recall value of 100% indicates that no pre-ictal segment was misclassified. Precision measures the proportion of true pre-ictal segments among all segments predicted as pre-ictal and is defined as $Prec = [TP/(TP+FP)] \cdot 100\%$. A precision value of 100% indicates that no inter-ictal segment was misclassified.

The AUC is a scalar metric summarizing the performance of a binary classifier across all classification thresholds, by relating the false positive rate (FPR) against true positive rate (TPR), forming the ROC. The closer the ROC is to the top-left corner, the better the classifier performance. The area under this curve summarizes the relationship between FPR and TPR [13]. To compute the AUC, where n is the total number

TABLE I:

NUMBER OF SAMPLES PER SUBJECT.						
Subject	1	2	3	4	5	6
N ^o seizures	7	2	6	3	5	5
N ^o total samples	8400	2400	7200	3600	6000	6000

of ROC points, the following trapezoidal rule can be used

$$AUC = \sum_{i=1}^{n-1} \frac{TPR_{i+1} + TPR_i}{2} \cdot (FPR_{i+1} - FPR_i). \quad (7)$$

C. Methodology

We applied A1 and A2 to classify inter- and pre-ictal phases. For each 3-second EEG segment, corresponding to 768 time-domain samples (T), the CWT was computed using 80 frequency bins (F) and 21 channels (Ch), forming a 3rd order spectral-temporal-spatial tensor segment.

Next, the SECSI framework proposed in [8] was used to compute the CPD of the EEG tensor segments. SECSI provides a semi-closed-form solution based on simultaneous matrix diagonalization (SMDs), avoiding the iterative optimization required by alternating least squares (ALS) and offering a favorable complexity-accuracy trade-off through different heuristic criteria for selecting the final factor-matrix estimates. In this work, the reconstructed paired solutions (REC PS) heuristic was adopted to select the appropriate SMD, since it provides robust performance with lower computational complexity when compared with CPD via ALS [8]. To evaluate different data-compression levels, the rank R was varied from 1 to 5, since EEG tensors exhibit a low-rank structure, as reported in [2].

In this work, a linear-kernel SVM was used to classify the feature vectors obtained with A1 and A2, as described in Section III. The box constraint and kernel scale were both set to 1, and the sequential minimal optimization (SMO) solver was used to handle the resulting large datasets efficiently.

The SVM performance was evaluated using a patient-specific LOSOCV scheme, where each CV fold corresponds to one seizure event. Recall, precision, and AUC were analyzed as functions of the number of accumulated features, which were reordered according to their FSs. Due to the structural differences between A1 and A2, distinct feature-evaluation step sizes were used. For A1, the evaluation was performed in rank-dependent increments of 100 features, since its dimensionality scales with $R(F+Ch)$. For A2, the evaluation was performed in rank-independent increments of 84 features, up to the same total feature ceiling of 8400. To reduce the influence of outliers associated with intra-patient seizure variability, the median performance across all cross-validation iterations is reported.

D. Fisher score analysis of CPD-derived features

Due to the high dimensionality of the CPD-based feature space and the number of samples (see Table I), FS analysis is used to assess feature separability for the binary classification task without relying on visual inspection of feature distributions (see Section IV-B). Fig. 7 depicts the FSs associated with the frequency (green), time (yellow), and channel (purple) domains for the rank-one CPD estimates obtained for each EEG band, $\nu \in \{\delta, \theta, \alpha, \beta, \gamma\}$. The jitter plot shows the FSs normalized by the maximum value across all subjects, bands, and rank-one components, whereas the pie charts show the average contribution of each domain. The corresponding number of features is given by $R(F+T+Ch)$.

Time-domain features consistently provide the lowest discriminative contribution (0.02–0.05) across the selected ranks

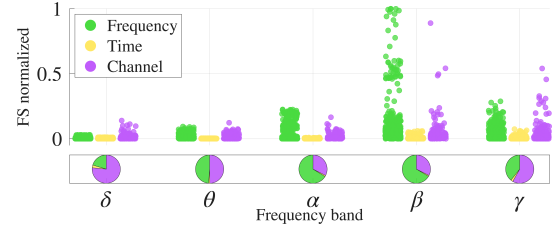


Fig. 7. FSs of CPD-derived features for $R = 1$, across the six subjects for the EEG bands δ , θ , α , β , and γ , including all components from A^ν , B^ν and C^ν . N^o features = 869.

($R = 1, \dots, 5$). In contrast, spatial features provide the largest contribution (0.6–0.8), becoming increasingly dominant for $R = 3, \dots, 5$, which suggests that inter-channel relationships play an important role in distinguishing inter- and pre-ictal phases. These results indicate a dominance of the spatial and spectral domains across all frequency bands. In particular, higher selected ranks emphasize spatial discriminative patterns, whereas temporal contributions remain limited and spectral features are mainly captured by lower-rank components.

E. Results

For both approaches, the rank R was varied from 1 to 5 to evaluate the classification performance between inter- and pre-ictal phases. Fig. 8 shows the average performance across subjects for each rank R as a function of the number of selected Fisher features (FFs). While both approaches achieve classification metrics above 70%, A1 requires significantly fewer features than A2 to reach competitive classification performance. For example, at $R = 5$, A1 uses 2525 features and improves precision and AUC by approximately 0.84 compared with A2, which uses 8400 features. Thus, A1 matches or outperforms A2 using approximately 3.3 $\times$ fewer features.

Except for $R = 1$, the performance of A1 improves with the rank R , narrowing the gap with A2 as more features are included in the SVM. Although A1 achieves satisfactory performance with a smaller feature set, A2 shows greater stability across different numbers of selected FFs. Moreover, A2 reaches performance saturation as features are accumulated, even under the highest compression level at $R = 1$.

Since the $R = 5$ configuration yields the best overall performance, with A1 benefiting from a larger feature set and A2 from a higher-resolution spectral-spatial representation, this configuration is further evaluated across subjects to assess performance consistency. Fig. 9 shows the best observed SVM performance, in terms of recall, precision, and AUC, achieved at $R = 5$ for each subject using a specific number of Fisher features (NFs).

The results show robust classification performance for subjects 1, 3, and 6, with recall and AUC above 95%. For these subjects, both A1 and A2 perform well, although A1 often achieves comparable performance with fewer features. For instance, for subject 3, A1 achieves 96.7% recall using only 504 features, closely matching the 96.8% recall obtained by A2 using 1764 features.

Subjects 2, 4, and 5 are more challenging, with recall values decreasing toward 70–80%. For subjects 2 and 4, this behavior may be related to the limited number of training samples, whereas for subject 5 it suggests higher intra-subject variability or stronger overlap between the inter- and pre-ictal

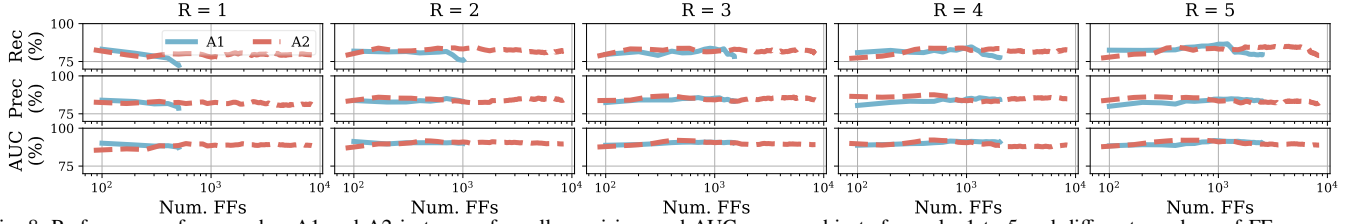


Fig. 8. Performance of approaches A1 and A2 in terms of recall, precision, and AUC across subjects for ranks 1 to 5 and different numbers of FFs.

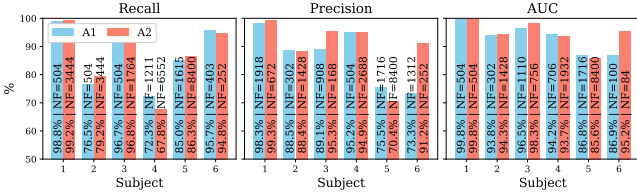


Fig. 9. Performance of approaches A1 and A2 in terms of best observed recall, precision, and AUC across all subjects for $R=5$.

feature distributions. These factors can limit the stability of the classification performance. To enhance classification stability and pre-ictal separability, investigations on incorporating additional seizure data, adaptive classification strategies, and alternative tensor decompositions can be employed.

The efficiency gains of A1 are subject-dependent. In several cases, A1 achieves comparable or superior results with a smaller NF, whereas A2 may require substantially more features to reach its best performance. For example, in the more challenging cases, the best recall, precision, and AUC for A2 require 3444, 1428, and 1428 features for subject 2; 6552, 2688, and 1932 features for subject 4; and the full 8400 features for subject 5. This indicates that the reduced compression of A2 does not necessarily lead to superior classification performance.

Across subjects, precision is more stable than recall, indicating that the classifier effectively minimizes false pre-ictal detections. However, recall, which measures the ability to identify all pre-ictal segments, remains more sensitive to patient-specific variability in seizure-related EEG patterns.

The AUC performance gap between A1 and A2 is small in most cases. Even when recall or precision decreases for a given subject, the AUC remains above 93% for all subjects except subject 5, indicating that the SVM preserves good overall discrimination capability. This suggests that FS-based feature selection provides informative features for separating inter- and pre-ictal segments, although the operating point may vary with the selected feature subset and subject-specific characteristics.

For challenging subjects (2, 4, and 5), A2 requires high feature counts (NFs) for peak metrics, demanding up to 3444, 6552, and the full 8400 features, respectively.

Overall, A1 tends to achieve comparable or superior results under higher compression, particularly for challenging subjects, while A2 may provide marginal gains in specific metrics at the cost of an NF that is approximately 3 to 5 times higher.

Compared with recent patient-specific CHB-MIT methods averaged across the same 6 subjects used in this work, A1 and A2 achieved competitive AUCs (93.00% and 94.50%) with recalls of 87.50% and 87.30%. While [5]–[7] report higher recalls (91.67%–98.98%), A2 maintains a comparable AUC of 94.50% to [7] (95.20%).

V. CONCLUSIONS

In this work, the use of a CPD to extract interpretable features from EEG tensors for epileptic seizure detection was investigated. Two approaches for inter- and pre-ictal classification, namely A1 and A2, were presented and evaluated. Both approaches achieved competitive classification performance; however, A1 matched or outperformed A2 while using fewer features. Furthermore, A2 showed that combining the estimated factor matrices to construct spectral-spatial signatures can provide stable classification performance. For the future work, models based on Fuzzy logic will be explored using features derived from the proposed approaches to identify intermediate and interpretable transitions among seizure phases.

REFERENCES

- [1] F. Mormann, T. Kreuz, *et al.*, “On the predictability of epileptic seizures,” *Clinical Neurophysiology*, vol. 116, no. 3, pp. 569–587, Mar. 2005, issn: 1388-2457. doi: 10.1016/j.clinph.2004.08.025.
- [2] Y. R. Aldana, B. Hunyadi, *et al.*, “Nonconvulsive epileptic seizure detection in scalp EEG using multiway data analysis,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 660–671, Mar. 2019, issn: 2168-2208. doi: 10.1109/jbhi.2018.2829877.
- [3] L. T. Thanh, N. T. A. Dao, *et al.*, “Multi-channel EEG epileptic spike detection by a new method of tensor decomposition,” *Journal of Neural Engineering*, vol. 17, no. 1, p. 016023, Jan. 2020, issn: 1741-2552. doi: 10.1088/1741-2552/ab5247.
- [4] T. Gupta, N. Wagh, *et al.*, “Tensor decomposition of large-scale clinical EEGs reveals interpretable patterns of brain physiology,” in *2023 11th International IEEE/EMBS Conference on Neural Engineering (NER)*, IEEE, 2023, pp. 1–4.
- [5] P. Koutsouvelis, B. Chybowski, *et al.*, “Preictal period optimization for deep learning-based epileptic seizure prediction,” *Journal of Neural Engineering*, vol. 21, no. 6, p. 066040, Dec. 2024, issn: 1741-2552. doi: 10.1088/1741-2552/ad9ad0.
- [6] J. Xiang, Y. Li, *et al.*, “Synchronization-based graph spatio-temporal attention network for seizure prediction,” *Scientific Reports*, vol. 15, no. 1, Feb. 2025, issn: 2045-2322. doi: 10.1038/s41598-025-88492-5.
- [7] L. Meng, L. Zhou, *et al.*, “Robust epileptic seizure prediction: A 3d-seresnet framework for patient-specific and multi-patient generalization,” *iScience*, vol. 28, no. 12, p. 114171, Dec. 2025, issn: 2589-0042. doi: 10.1016/j.isci.2025.114171.
- [8] F. Roemer and M. Haardt, “A semi-algebraic framework for approximate CP decompositions via simultaneous matrix diagonalizations (SECSI),” *Signal Processing*, vol. 93, no. 9, pp. 2722–2738, Sep. 2013, issn: 0165-1684. doi: 10.1016/j.sigpro.2013.02.016.
- [9] M. Steriade, P. Gloor, *et al.*, “Basic mechanisms of cerebral rhythmic activities,” *Electroencephalography and Clinical Neurophysiology*, vol. 76, no. 6, pp. 481–508, Dec. 1990, issn: 0013-4694. doi: 10.1016/0013-4694(90)90001-z.
- [10] R. A. Harshman *et al.*, “Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis,” *UCLA working papers in phonetics*, vol. 16, no. 1, p. 84, 1970.
- [11] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd. Wiley Hoboken, 2001, isbn: 978-0-471-05669-0.
- [12] A. Shoeb, *CHB-MIT scalp EEG database*, 2010. doi: 10.13026/C2K01R.
- [13] J. A. Hanley and B. J. McNeil, “The meaning and use of the area under a receiver operating characteristic (ROC) curve,” *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.