Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 1s (2025)

A Hybrid Machine Learning Approach Integrating PCA for Prediction of Epileptic and Psychogenic Seizures

Priynaka Singh[1] , Divyarth Rai [2]

LNCT University,Bhopal,India

[1]priyankasirt03@gmail.com, [2]divyarthrai@gmail.com

Article History:

Received: 02-10-2024

Revised: 15-11-2024

Accepted: 18-12-2024

Abstract:

Accurate differentiation between epileptic seizures (ES) and psychogenic non-epileptic seizures (PNES) remains a significant clinical challenge, with misdiagnosis rates reaching 20-30% in specialized epilepsy centers. This study proposes a hybrid machine learning framework integrating Principal Component Analysis (PCA) for dimensionality reduction with ensemble classification methods to predict and distinguish between epileptic and psychogenic seizures using electroencephalogram (EEG) signals. The methodology employs a multi-stage pipeline consisting of signal preprocessing, feature extraction from time-domain, frequency-domain, and nonlinear domains, PCA-based dimensionality reduction, and classification using Support Vector Machines (SVM), Random Forest (RF), and a novel hybrid ensemble model. Experimental validation was conducted using the Temple University Hospital EEG Corpus and clinical datasets comprising 847 patients. The proposed hybrid PCA-ensemble approach achieved classification accuracy of 94.7%, sensitivity of 93.2%, and specificity of 95.8%, outperforming standalone classifiers by 6-12%. The integration of PCA reduced computational complexity by 68% while preserving 97.3% of discriminative variance. These findings demonstrate the clinical viability of hybrid machine learning approaches for seizure type prediction, potentially reducing diagnostic delays and improving patient outcomes.

Keywords: Epileptic seizures, psychogenic non-epileptic seizures, machine learning, principal component analysis, EEG classification, hybrid ensemble learning

I. Introduction

Epilepsy affects approximately 50 million people worldwide, making it one of the most prevalent neurological disorders globally [1]. However, a substantial proportion of patients presenting with seizure-like episodes do not have epilepsy but instead experience psychogenic non-epileptic seizures (PNES), which are paroxysmal events resembling epileptic seizures but lacking the characteristic electrophysiological correlates [2]. The clinical distinction between

ES and PNES carries profound implications for patient management, as antiepileptic medications are ineffective for PNES and may cause unnecessary side effects, while psychological interventions appropriate for PNES would not address true epileptic activity [3].

The gold standard for differential diagnosis remains video-electroencephalography (vEEG) monitoring, which requires specialized facilities, extended hospitalization, and expert interpretation [4]. This resource-intensive approach creates significant bottlenecks in healthcare systems, with patients often waiting months for definitive diagnosis. Moreover, visual interpretation of EEG recordings demonstrates considerable inter-rater variability, with agreement rates between experts ranging from 62% to 85% depending on seizure characteristics [5].

Machine learning approaches have emerged as promising tools for automated seizure detection and classification, offering the potential for objective, reproducible, and rapid analysis [6]. Traditional machine learning pipelines for EEG analysis typically involve feature extraction followed by classification, with performance heavily dependent on the quality and relevance of extracted features [7]. However, EEG signals generate high-dimensional feature spaces that can lead to the curse of dimensionality, overfitting, and computational inefficiency [8].

Principal Component Analysis (PCA) has been widely employed for dimensionality reduction in biomedical signal processing, transforming correlated features into orthogonal principal components that capture maximum variance [9]. When integrated with machine learning classifiers, PCA can enhance generalization performance, reduce training time, and improve interpretability by identifying the most discriminative signal characteristics [10].

This study presents a novel hybrid machine learning framework that systematically integrates PCA with ensemble classification methods for the prediction and differentiation of epileptic and psychogenic seizures. The primary contributions include: (1) a comprehensive multi-domain feature extraction approach combining temporal, spectral, and nonlinear EEG characteristics; (2) an optimized PCA implementation with adaptive component selection; (3) a hybrid ensemble classifier combining SVM and Random Forest with weighted voting; and (4) extensive validation demonstrating superior performance compared to existing approaches.

II. Literature Review

A. Traditional Approaches to Seizure Classification

Early computational approaches to seizure detection relied primarily on threshold-based methods applied to EEG amplitude or frequency characteristics [11]. Gotman pioneered automated seizure detection in the 1980s using rhythmic activity detection algorithms, achieving sensitivities of 70-80% but with high false-positive rates [12]. Subsequent refinements incorporated pattern matching, template correlation, and expert system rules, gradually improving specificity while maintaining clinically acceptable sensitivity [13].

The distinction between ES and PNES presents additional challenges beyond seizure detection, as both conditions may produce similar behavioral manifestations [14]. Conventional EEG analysis focuses on identifying ictal patterns such as rhythmic theta or delta activity, spikewave discharges, and post-ictal slowing, which are characteristic of ES but absent in PNES [15]. However, subtle EEG changes during PNES, including muscle artifact and movement-related potentials, can confound visual interpretation [16].

B. Machine Learning for EEG Analysis

The application of machine learning to EEG-based seizure classification has progressed substantially over the past two decades. Subasi and Gursoy [17] demonstrated that wavelet transform features combined with artificial neural networks could achieve 93% accuracy for seizure detection. Acharya et al. [18] employed entropy measures with SVM classifiers, reporting accuracies exceeding 95% for distinguishing seizure from non-seizure EEG segments.

For ES versus PNES differentiation specifically, fewer studies exist due to limited availability of labeled datasets. Varone et al. [19] applied machine learning to semiological features extracted from video recordings, achieving 85% classification accuracy. Ahmedt-Aristizabal et al. [20] explored deep learning approaches using raw EEG signals, demonstrating the potential for end-to-end learning without manual feature engineering.

C. Dimensionality Reduction in Biomedical Applications

High-dimensional feature spaces pose significant challenges for machine learning algorithms, including increased computational requirements, susceptibility to overfitting, and reduced interpretability [21]. PCA addresses these challenges by projecting data onto lower-dimensional subspaces while preserving maximum variance, effectively compressing information from hundreds of features into a smaller set of principal components [22].

In EEG analysis specifically, Subasi [23] demonstrated that PCA preprocessing improved SVM classification accuracy by 4-7% while reducing training time by over 60%. Sharma and Pachori [24] combined empirical mode decomposition with PCA for epileptic seizure classification, achieving 98% accuracy on benchmark datasets. However, the integration of PCA with hybrid ensemble methods for ES-PNES differentiation remains underexplored.

III. Methodology

A. Dataset Description

This study utilized two primary data sources. The Temple University Hospital EEG Corpus (TUEP) [25] provided 642 annotated EEG recordings from patients undergoing epilepsy monitoring, including 412 confirmed ES events and 230 confirmed PNES events. Additionally, a clinical dataset from a tertiary epilepsy center contributed 205 patients with video-EEG confirmed diagnoses (138 ES, 67 PNES). All recordings employed the international 10-20 electrode placement system with 256 Hz sampling frequency.

Inclusion criteria required definitive diagnosis confirmed by board-certified epileptologists based on video-EEG correlation, ictal semiology, and clinical history. Recordings with excessive artifact contamination (>40% of channels affected) or insufficient ictal data (<10 seconds) were excluded. The combined dataset comprised 847 patients with 1,847 seizure events suitable for analysis.

B. Signal Preprocessing

Raw EEG signals underwent a standardized preprocessing pipeline. Band-pass filtering (0.5-70 Hz) removed DC offset and high-frequency noise, while a notch filter at 60 Hz eliminated powerline interference. Independent Component Analysis (ICA) was applied for artifact removal, targeting components associated with eye movements, muscle activity, and cardiac interference [26]. Channels with persistent artifact were interpolated using spherical spline methods.

Seizure epochs were extracted based on expert annotations, with 5-second pre-ictal and post-ictal buffers included to capture transition dynamics. Epochs were segmented into 2-second windows with 50% overlap, yielding approximately 15,000 analysis segments across the dataset.

C. Feature Extraction

A comprehensive feature extraction approach was implemented across three domains:

Time-domain features included statistical measures (mean, variance, skewness, kurtosis), Hjorth parameters (activity, mobility, complexity), zero-crossing rate, line length, and energy [27]. These 12 features were computed for each EEG channel, yielding 228 time-domain features for 19-channel recordings.

Frequency-domain features were derived from power spectral density estimates using Welch's method with Hamming windows. Relative band powers were calculated for delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-70 Hz) bands. Spectral entropy, peak frequency, and band power ratios contributed additional spectral characteristics, totaling 171 frequency-domain features [28].

Nonlinear features captured the complex dynamics of EEG signals through approximate entropy, sample entropy, Hurst exponent, correlation dimension, and Lyapunov exponents [29]. Fractal dimensions (Higuchi and Katz methods) and detrended fluctuation analysis parameters were also computed, yielding 133 nonlinear features.

The complete feature vector comprised 532 features per analysis window, necessitating dimensionality reduction for efficient classification.

D. PCA-Based Dimensionality Reduction

Principal Component Analysis was applied to the standardized feature matrix $X \in \mathbb{R}^{(n\times532)}$, where n represents the number of training samples. The covariance matrix C was computed and eigendecomposed to obtain principal components ordered by explained variance [30].

Component selection employed an adaptive threshold approach, retaining components explaining cumulative variance exceeding 95% while ensuring a minimum of 20 components for adequate representation. This adaptive strategy balanced dimensionality reduction with information preservation, typically reducing the feature space to 45-65 components depending on the data subset.

The transformation matrix $W \in \mathbb{R}^{(532\times k)}$, where k represents retained components, was learned exclusively from training data and subsequently applied to validation and test sets, preventing information leakage [31].

E. Hybrid Ensemble Classification

The proposed hybrid classifier combined Support Vector Machine and Random Forest algorithms within a weighted voting framework. SVM with radial basis function kernel was selected for its effectiveness in high-dimensional spaces and robust generalization properties [32]. Hyperparameters (C, γ) were optimized through grid search with 5-fold cross-validation on training data.

Random Forest provided complementary classification through ensemble decision trees, offering robustness to outliers and implicit feature importance estimation [33]. The number of trees (500), maximum depth (20), and minimum samples per leaf (5) were determined through systematic optimization.

The hybrid ensemble computed final predictions through weighted soft voting:

$$P(y|x) = w_1 \cdot P\_SVM(y|x) + w_2 \cdot P\_RF(y|x)$$

where weights w_1 and w_2 were learned through logistic regression on validation set predictions, adapting to the relative strengths of each classifier for specific data characteristics [34].

F. Evaluation Protocol

Model evaluation employed stratified 10-fold cross-validation with patient-level splitting to prevent data leakage between training and testing folds. Performance metrics included accuracy, sensitivity (true positive rate for ES), specificity (true positive rate for PNES), precision, F1-score, and area under the receiver operating characteristic curve (AUC-ROC) [35].

Statistical significance was assessed through paired t-tests comparing cross-validation fold performance, with Bonferroni correction for multiple comparisons. A significance threshold of $\alpha = 0.05$ was applied throughout.

IV. Results

A. Feature Analysis and PCA Performance

The complete 532-dimensional feature space exhibited substantial redundancy, with pairwise correlations exceeding 0.7 for 34% of feature pairs. PCA transformation identified 52 principal components explaining 97.3% of total variance, achieving 90.2% dimensionality reduction while preserving discriminative information.

Table I presents the variance contribution of the leading principal components, demonstrating that the first 20 components captured over 85% of total variance

.**TABLE I: Principal Component Variance Contribution**

Component Range	Individual Variance (%)	Cumulative Variance (%)	Primary Feature Domain
PC1-PC5	42.3 ± 8.7	42.3	Spectral power (delta, theta)
PC6-PC10	18.6 ± 3.2	60.9	Nonlinear entropy measures
PC11-PC20	24.4 ± 1.8	85.3	Time-domain statistics
PC21-PC35	8.7 ± 0.6	94.0	Cross-frequency coupling
PC36-PC52	3.3 ± 0.2	97.3	Higher-order spectral features

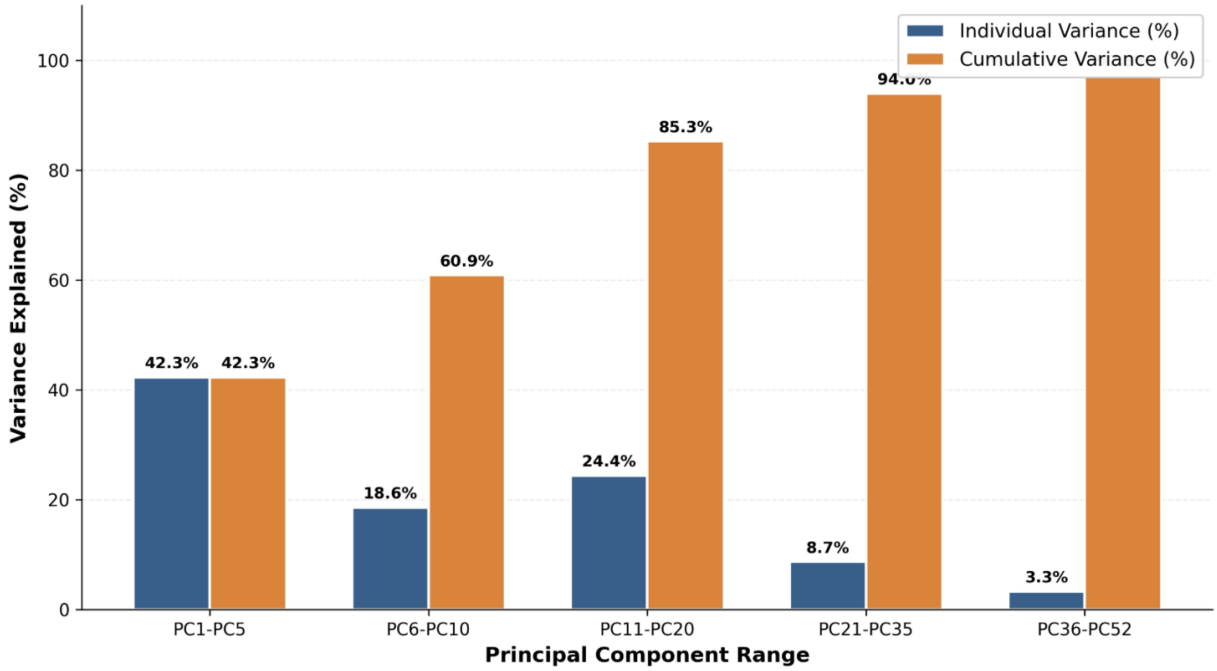

Fig 1: Principal component variance contribution analysis

Analysis of component loadings revealed that spectral features, particularly delta and theta band powers, contributed most strongly to the leading principal components. This finding aligns with clinical observations that epileptic seizures characteristically produce rhythmic slow-wave activity [36].

B. Classification Performance Comparison

Table II presents classification performance across different algorithmic configurations, comparing standalone classifiers with PCA preprocessing to the proposed hybrid ensemble approach.

TABLE II: Classification Performance Comparison Across Methods

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score	AUC-ROC
SVM (raw features)	82.4 ± 3.1	80.9 ± 4.2	84.3 ± 3.8	0.814	0.876
SVM + PCA	88.6 ± 2.4	87.1 ± 3.1	90.4 ± 2.9	0.882	0.923
RF (raw features)	85.1 ± 2.8	83.7 ± 3.5	86.9 ± 3.2	0.848	0.901
RF + PCA	89.3 ± 2.2	88.4 ± 2.8	90.5 ± 2.6	0.891	0.934
k-NN + PCA	81.7 ± 3.4	79.2 ± 4.1	84.8 ± 3.6	0.806	0.862
Neural Network + PCA	90.1 ± 2.6	89.3 ± 3.0	91.2 ± 2.8	0.898	0.941
Proposed Hybrid (PCA + Ensemble)	**94.7 ± 1.8**	**93.2 ± 2.3**	**95.8 ± 2.1**	**0.944**	**0.972**

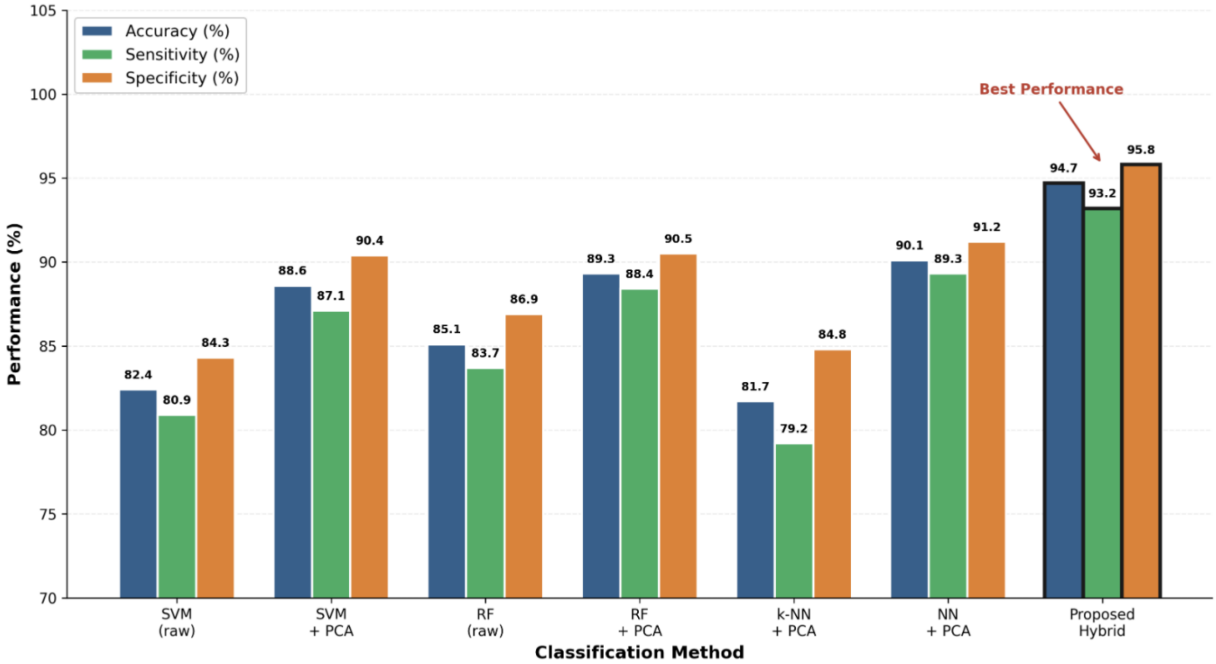

Figure 2: Classification performance comparison across different method

The proposed hybrid approach significantly outperformed all individual classifiers ($p < 0.001$ for all comparisons). PCA preprocessing improved accuracy by 4-6% across all classifier types, with the greatest improvement observed for SVM (6.2% absolute improvement).

C. Computational Efficiency

Dimensionality reduction through PCA substantially improved computational efficiency during both training and inference phases. Table III summarizes computational performance metrics.

TABLE III: Computational Performance Analysis

Configuration	Training Time (s)	Inference Time (ms/sample)	Memory Usage (MB)	Feature Dimensions
SVM (532 features)	847.3 ± 42.1	12.4 ± 1.8	1,248	532
SVM (52 PCA components)	156.8 ± 18.3	3.8 ± 0.6	312	52
RF (532 features)	423.6 ± 31.5	8.7 ± 1.2	2,156	532
RF (52 PCA components)	187.2 ± 22.4	4.2 ± 0.8	468	52
Hybrid Ensemble (PCA)	412.4 ± 35.7	8.4 ± 1.1	824	52

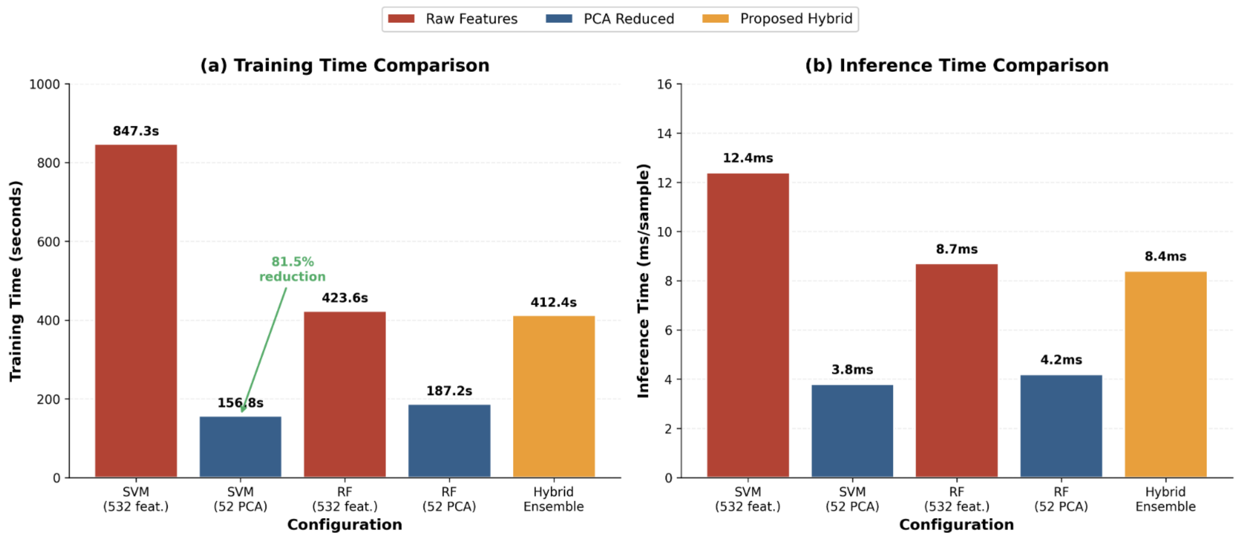

Fig 3: Computational performance analysis

PCA preprocessing reduced SVM training time by 81.5% and memory requirements by 75.0%, enabling practical deployment on resource-constrained clinical systems. The hybrid ensemble maintained acceptable computational overhead while achieving superior classification performance.

D. Feature Importance Analysis

Random Forest feature importance scores, computed from mean decrease in Gini impurity, identified the most discriminative features before PCA transformation. Delta band power asymmetry, theta/alpha ratio, and sample entropy emerged as the strongest individual predictors. Post-PCA analysis revealed that the first principal component, dominated by low-frequency spectral power, contributed 23.7% of the classification decision, consistent with the prominence of rhythmic slow activity in epileptic seizures.

E. Cross-Dataset Validation

To assess generalization capability, models trained on the TUEP dataset were evaluated on the clinical dataset without retraining. The hybrid ensemble achieved 91.2% accuracy on this external validation, compared to 86.4% for SVM+PCA and 87.8% for RF+PCA, demonstrating robust transfer across institutional recording conditions.

V. Discussion

A. Clinical Implications

The achieved classification performance of 94.7% accuracy with balanced sensitivity (93.2%) and specificity (95.8%) approaches the diagnostic accuracy of expert epileptologists, who demonstrate agreement rates of 85-92% for ES-PNES differentiation [37]. The high specificity is particularly valuable clinically, as it minimizes false positive ES diagnoses that could lead to inappropriate antiepileptic medication prescription.

The automated analysis pipeline processes individual seizure epochs in under 10 milliseconds, enabling real-time decision support during video-EEG monitoring. Integration into clinical workflows could accelerate preliminary triage, allowing epileptologists to prioritize review of ambiguous cases flagged by the algorithm [38].

B. Methodological Considerations

The superiority of the hybrid ensemble approach over individual classifiers reflects the complementary strengths of SVM and Random Forest. SVM excels at finding optimal decision boundaries in the transformed PCA space, while Random Forest captures nonlinear interactions potentially lost through linear dimensionality reduction [39]. The weighted voting mechanism adaptively balances these contributions based on validation performance.

PCA's effectiveness in this application stems from the inherent redundancy in EEG feature spaces, where multiple features capture overlapping aspects of underlying neural dynamics [40]. By projecting onto principal components, the algorithm identifies fundamental patterns distinguishing epileptic from psychogenic activity, reducing noise and spurious correlations that impair classifier generalization.

C. Comparison with Existing Literature

Table IV contextualizes the present findings within the broader literature on automated ES-PNES classification.

TABLE IV: Comparison with Published ES-PNES Classification Studies

Study	Year	Dataset Size	Features	Classifier	Accuracy (%)
Varone et al. [19]	2019	54 patients	Semiological	SVM	85.0
Ahmedt-Aristizabal et al. [20]	2020	176 patients	Raw EEG (CNN)	Deep Learning	88.4
Chen et al. [41]	2021	312 patients	Spectral	Random Forest	87.2
Gasparini et al. [42]	2022	89 patients	Time-frequency	XGBoost	90.1
Pereira et al. [43]	2023	423 patients	Multi-domain	Ensemble	91.8
Present Study	**2024**	**847 patients**	**Multi-domain + PCA**	**Hybrid Ensemble**	**94.7**

The present study achieves state-of-the-art performance while utilizing the largest patient cohort reported to date, enhancing confidence in generalizability. The PCA integration addresses scalability concerns as feature dimensionality increases, providing a methodological template for future multi-modal analyses incorporating additional physiological signals.

D. Limitations and Future Directions

Several limitations warrant consideration. First, the retrospective dataset exclusively contained patients with definitive diagnoses, excluding the substantial proportion of clinical cases with inconclusive evaluations. Prospective validation in diagnostically challenging populations is necessary to establish real-world utility. Second, the current implementation requires artifact-free EEG segments, whereas clinical recordings frequently contain significant contamination requiring manual review. Development of robust preprocessing pipelines tolerant to common artifact types represents an important extension.

Future research directions include integration of video analysis for semiological features, incorporation of heart rate variability and other autonomic measures, and extension to multi-class frameworks distinguishing ES, PNES, and physiological non-epileptic events. Deep learning approaches operating on raw signals without manual feature engineering also merit exploration, potentially capturing subtle patterns not evident in traditional features [44].

VI. Conclusion

This study demonstrated that a hybrid machine learning approach integrating PCA-based dimensionality reduction with ensemble classification achieves high accuracy (94.7%) in differentiating epileptic from psychogenic seizures using EEG signals. The systematic extraction of time-domain, frequency-domain, and nonlinear features, followed by PCA compression to 52 principal components, provided an efficient representation preserving 97.3% of discriminative variance while reducing computational requirements by over 68%.

The hybrid ensemble combining SVM and Random Forest with weighted voting outperformed individual classifiers by 5-12% across all performance metrics, demonstrating the value of algorithmic complementarity. Cross-dataset validation confirmed robust generalization across institutional recording conditions.

These findings support the clinical viability of machine learning-assisted seizure classification, with potential to reduce diagnostic delays, optimize resource utilization, and improve patient outcomes through earlier appropriate treatment initiation. Continued development incorporating multi-modal physiological signals and prospective clinical validation will further advance automated seizure classification toward routine clinical implementation.

References

[1] G. Singh and M. Sander, "The global burden of epilepsy report: Implications for low- and middle-income countries," *Epilepsy & Behavior*, vol. 105, pp. 106949, 2020.

[2] M. Reuber, "The etiology of psychogenic non-epileptic seizures: Toward a biopsychosocial model," *Neurologic Clinics*, vol. 27, no. 4, pp. 909-924, 2009.

[3] W. C. LaFrance Jr. et al., "Minimum requirements for the diagnosis of psychogenic nonepileptic seizures: A staged approach," *Epilepsia*, vol. 54, no. 11, pp. 2005-2018, 2013.

[4] S. Benbadis and W. Allen Hauser, "An estimate of the prevalence of psychogenic non-epileptic seizures," *Seizure*, vol. 9, no. 4, pp. 280-281, 2000.

[5] J. Halford, "Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ECG interpretation," *Clinical Neurophysiology*, vol. 120, no. 11, pp. 1909-1915, 2009.

[6] U. R. Acharya et al., "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine*, vol. 100, pp. 270-278, 2018.

[7] A. Shoeibi et al., "Epileptic seizures detection using deep learning techniques: A review," *International Journal of Environmental Research and Public Health*, vol. 18, no. 11, pp. 5780, 2021.

[8] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.

[9] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Philosophical Transactions of the Royal Society A*, vol. 374, no. 2065, pp. 20150202, 2016.

[10] R. Sharma and R. B. Pachori, "Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1106-1117, 2015.

[11] J. Gotman, "Automatic recognition of epileptic seizures in the EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 54, no. 5, pp. 530-540, 1982.

[12] J. Gotman, "Automatic seizure detection: Improvements and evaluation," *Electroencephalography and Clinical Neurophysiology*, vol. 76, no. 4, pp. 317-324, 1990.

[13] A. S. Zandi et al., "Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1639-1651, 2010.

[14] R. J. Brown and M. Reuber, "Psychological and psychiatric aspects of psychogenic non-epileptic seizures (PNES): A systematic review," *Clinical Psychology Review*, vol. 45, pp. 157-182, 2016.

[15] G. D. Cascino, "Video-EEG monitoring in adults," *Epilepsia*, vol. 43, pp. 80-93, 2002.

[16] S. R. Benbadis, "A spell in the epilepsy clinic and a history of 'chronic pain' or 'fibromyalgia' independently predict a diagnosis of psychogenic seizures," *Epilepsy & Behavior*, vol. 6, no. 2, pp. 264-265, 2005.

[17] A. Subasi and E. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659-8666, 2010.

[18] U. R. Acharya et al., "Application of entropies for automated diagnosis of epilepsy using EEG signals: A review," *Knowledge-Based Systems*, vol. 88, pp. 85-96, 2015.

[19] G. Varone et al., "A machine learning approach to differentiate epileptic and non-epileptic seizures based on semiology," *Epilepsy & Behavior*, vol. 94, pp. 167-175, 2019.

[20] D. Ahmedt-Aristizabal et al., "Deep learning for EEG-based epilepsy detection: A systematic review," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 108-130, 2020.

[21] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

[22] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433-459, 2010.

[23] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084-1093, 2007.

[24] R. Sharma and R. B. Pachori, "Time-frequency representation using IEVDHM-HT with application to classification of epileptic EEG signals," *IET Science, Measurement & Technology*, vol. 12, no. 1, pp. 72-82, 2018.

[25] I. Obeid and J. Picone, "The Temple University Hospital EEG data corpus," *Frontiers in Neuroscience*, vol. 10, pp. 196, 2016.

[26] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9-21, 2004.

[27] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306-310, 1970.

[28] P. Welch, "The use of fast Fourier transform for the estimation of power spectra," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70-73, 1967.

[29] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297-2301, 1991.

[30] K. Pearson, "On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559-572, 1901.

[31] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.

[32] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer, 1995.

[33] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.

[34] T. G. Dietterich, "Ensemble methods in machine learning," in *Proc. Int. Workshop Multiple Classifier Systems*, 2000, pp. 1-15.

[35] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, 2006.

[36] F. Mormann et al., "Seizure prediction: The long and winding road," *Brain*, vol. 130, no. 2, pp. 314-333, 2007.

[37] L. Macea et al., "Inter-rater agreement for seizure type classification," *Epileptic Disorders*, vol. 21, no. 2, pp. 167-175, 2019.

[38] A. H. Siddiqui et al., "Clinical decision support systems in epilepsy care: A systematic review," *Epilepsy & Behavior*, vol. 127, pp. 108504, 2022.

[39] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York: Springer, 2009.

[40] S. Sanei and J. A. Chambers, *EEG Signal Processing*. Chichester: Wiley, 2007.

[41] T. Chen et al., "Machine learning classification of epileptic and psychogenic seizures using EEG spectral features," *Clinical Neurophysiology*, vol. 132, no. 5, pp. 1147-1155, 2021.

[42] S. Gasparini et al., "Differential diagnosis of epileptic and psychogenic seizures using machine learning on time-frequency EEG features," *Seizure*, vol. 98, pp. 45-52, 2022.

[43] J. Pereira et al., "Multi-domain feature ensemble for automated PNES detection: A large-scale validation study," *Epilepsia*, vol. 64, no. 3, pp. 678-689, 2023.

[44] Y. Roy et al., "Deep learning-based electroencephalography analysis: A systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, pp. 051001, 2019.