J Med Syst (2017) 41: 160
DOI 10.1007/s10916-017-0800-x

IMAGE & SIGNAL PROCESSING

Classification of Focal and Non Focal Epileptic Seizures Using Multi-Features and SVM Classifier

N. Sriraam[1] · **S. Raghu**[1]

Received: 13 July 2017 / Accepted: 14 August 2017 / Published online: 2 September 2017

Abstract Identifying epileptogenic zones prior to surgery is an essential and crucial step in treating patients having pharmacoresistant focal epilepsy. Electroencephalogram (EEG) is a significant measurement benchmark to assess patients suffering from epilepsy. This paper investigates the application of multi-features derived from different domains to recognize the focal and non focal epileptic seizures obtained from pharmacoresistant focal epilepsy patients from Bern Barcelona database. From the dataset, five different classification tasks were formed. Total 26 features were extracted from focal and non focal EEG. Significant features were selected using Wilcoxon rank sum test by setting p-value ($p < 0.05$) and z-score ($-1.96 > z > 1.96$) at 95% significance interval. Hypothesis was made that the effect of removing outliers improves the classification accuracy. Turkey's range test was adopted for pruning outliers from feature set. Finally, 21 features were classified using optimized support vector machine (SVM) classifier with 10-fold cross validation. Bayesian optimization technique was adopted to minimize the cross-validation loss. From the simulation results, it was inferred that the highest sensitivity, specificity, and classification accuracy of 94.56%, 89.74%, and 92.15% achieved respectively and found to be better than the state-of-the-art approaches. Further, it was observed that the classification accuracy improved from 80.2% with outliers to 92.15% without outliers. The classifier performance metrics ensures the suitability of the proposed multi-features with optimized SVM classifier. It can be concluded that the proposed approach can be applied for recognition of focal EEG signals to localize epileptogenic zones.

Keywords Classifier · EEG · Epileptic seizures · Focal and non focal · Multi-feature · SVM

This article is part of the Topical Collection on *Image & Signal Processing*

✉ N. Sriraam
sriraam@msrit.edu

S. Raghu
raghu.sidda@gmail.com

[1] Center for Medical Electronics and Computing, Ramaiah Institute of Technology (Affiliated to VTU Belgaum), Bangalore, 560054, India

Introduction

The temporary disruption in the neuronal brain region due to sudden burst is the indication of epileptic seizures. According to International League Against Epilepsy (ILAE) [12, 24], a standard topology of epilepsy classification is adopted by the neurological clinical community. Screening and interpretation of seizures is a tedious task for the clinicians due to the visual inspection procedure and they do rely on the computer aided qualitative tools for seizures recognition. In the last 10 decades, several breakthroughs have been attained for seizures detection related studies [3, 11, 33, 35]. Electroencephalogram (EEG) recordings merely carry the seizures activities which can be introspected quantitatively and qualitatively [12]. The area of brain cortex that initiates the seizures activities is referred as epileptogenic zones. The focal seizures also referred as complex partial seizures is predominant in the temporal lobe and also occurs in the frontal lobe for a shorter duration. Further, in certain conditions, a bilateral tonic-clonic seizure which lasts between 30 seconds and 3 minutes on both side of the brain cortex. Hence appropriate localization mechanism needs to

be adopted to quantify this region [25]. Focal EEG signals which are more non-linear and less random in nature are merely extracted from this epileptogenic zone and non focal EEG are recorded from another area of the epileptogenic zone [11]. The proposed study investigates the effect of outliers in multi-features extracted from focal and non focal EEG recordings followed by pattern classifier using optimized SVM classifier.

Literature review

Several attempts have been made to classify focal and non focal EEG signals in the past [4, 8, 11, 33–35]. Andrzejak et al. [4] have initiated an attempt using Bern Barcelona database for classification of focal and non focal activities. In this study, nonrandomness, nonlinear dependence, and nonstationarity of EEG signals were studied. In [33], focal and non focal EEG signals were decomposed up to six levels using discrete wavelet transform (DWT) to compute approximate and detail coefficients of sub-band signals. Average wavelet, permutation, fuzzy and phase entropies were computed from each sub-band followed by features ranking using Bhattacharyya space algorithm, Student's t-test, Wilcoxon test, receiver operating characteristic and entropy methods. Student's t-test and receiver operating characteristic-based ranking methods were found to be most suitable with highest average classification accuracy (CA) of 84%, sensitivity of 84% and specificity of 84% using the least square support vector machine (LS-SVM) classifier. It is reported in literature that different signal processing methods such as Fourier transform, DWT, wavelet packet transform, empirical mode decomposition (EMD) was used to decomposes EEG into various levels [8, 11, 33–35].

In study [11], EMD followed by DWT was applied on EEG signal to compute log energy entropy. They did not remove outliers from feature samples. The obtained feature were classified using K- nearest neighbor (K-NN) classifier which yield accuracy of 89.4%. Computational complexity was found to be high using EMD and DWT. Sharma et al. [34] extracted sample entropies and variances of the intrinsic mode functions obtained by EMD of EEG signals. Using LS-SVM classifier, accuracy of 85% was attained using two features. In another study [8], EEG signals were decomposed using various wavelet functions such as, bior5.5, coif3, db2, dmey, haar, rbio6.8, and sym6 to extract 9 different statistical features. The highest accuracy of 83% was achieved using SVM classifier with radial basis function as kernel.

Several automated recognition of epileptic seizures have been reported in past and provided the scope for introducing the procedure in clinical routine. Menshawy et al. [23] have used features mean, standard deviation, skewness, kurtosis, median in the first and second derivative of EEG signals. Bogaarts et al. [6] extracted features like curve length, root mean square, band power, zero crossing, Hjorth parameters (activity, mobility, and complexity), and Teager energy to classify epileptic EEG from normal. The effect of wavelet packet decomposition on EEG signals was studied using log energy entropy [29]. Different entropies like log energy, norm [30], wavelet, sample, and spectral [27], approximate [36], Renyi, Shannon [3] entropies were explored to classify between normal and epileptic EEG.

Li et al. [22] showed an improvement in CA from from 63% to 76% after removing outliers in multi-spectral burn diagnostic imaging [22]. Krauledat et al. [21] have applied outliers removal concept for classification of EEG signals. It was reported that outlier removal procedures can strongly enhance the classification performance. Similarly, Winkler et al. [37] employed hand-optimized selection of source components derived from independent component analysis to clean EEG data.

SVM optimization technique is capable of estimating excellent learning parameters, separating hyperplanes, and kernel functions. In [13], automatic approach to adjust the learning parameters of SVM was studied using a derivative free numerical optimizer. Optimal classification scheme was determined for EEG signal features derived from wavelet coefficients and Lyapounov exponents and it was reported high classification accuracies [18]. The features obtained from auto-regressive Burg power spectrum method was classified using optimal SVM classifier [28]. Regularization parameter and kernel parameter were calculated with a particle swarm optimization technique. Recently, optimization configuration of multi-layer perceptron neural network shown a good classification results for EEG signals [31]. It was understood from the studies [13, 18, 21, 22, 28, 37], the techniques, outliers removal and optimization of SVM classifier enhances the classification accuracy.

In the past, less CA was attained using a single feature which was not enough for clinical implementation. It was clear from the studies [6, 23], that the CA improves as the number of features increase. Taking this note into consideration, proposed method uses multi-features at the cost of computational complexity to achieve clinically acceptable results.

Proposed method

Figure 1 shows the step by step flow of proposed scheme. The database consists of focal and non focal EEG series with each having pair of signals. The paired EEG signals were named as "x" and "y" series. Totally, 26 features were extracted for 10 seconds segment. To find out significant features Wilcoxon rank sum test was performed. The outliers were excluded from the feature set using Turkey's range test and standardized prior to classification. In the

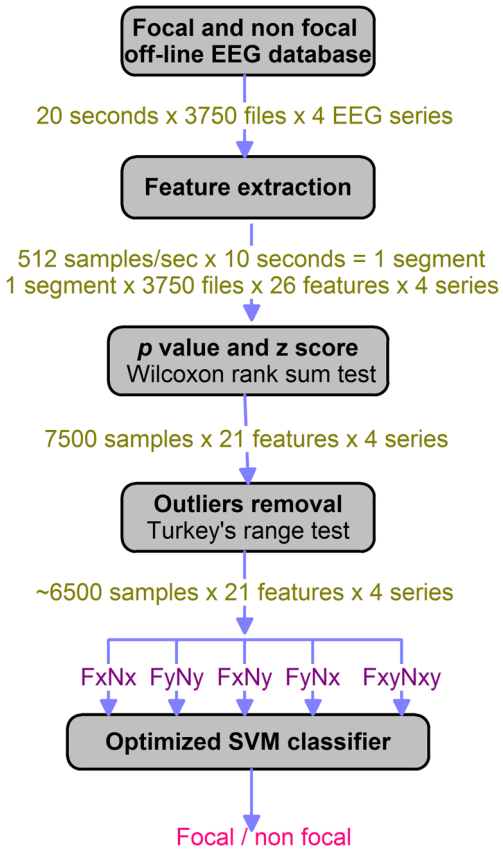

Fig. 1 Step by step flow of proposed focal and non focal seizures EEG classification

recent attempts on focal and non-focal classification problems, SVM classification found to be showing promising results. To enhance the classification efficiency with less false positive, the proposed work suggests a regularization parameter for multi features optimization. The final 21 features set were classified using SVM classifier with 10-fold cross validation. Bayesian optimization technique was adopted to minimize the cross-validation loss (error) for varying the parameters. Results were compared with and without outliers in features. Finally, obtained results were compared with state-of-the-art methods.

Materials and methods

EEG dataset

The focal and non focal EEG recording for this study was taken from open source Bern Barcelona database [4]. The motivation for us to use this database due to its huge volume and publicly available database. Apart from this, many researches have used same database for their studies [1, 8, 10, 33, 35]. This dataset includes intracranial EEG recordings from five pharmacoresistant epilepsy patients and electrodes were placed at the positions Fz and Pz. EEG signals were either sampled at 512 or 1024 Hz for more or less than 64 channels respectively. Later, EEG signals recorded with a sampling rate of 1024 Hz were down-sampled to 512 Hz. All EEG signals were passed through fourth order Butterworth band-pass filter (0.5 Hz-150 Hz). The database consists of 3750 focal and 3750 non focal EEG signal pairs of 20 seconds duration each. Focal EEG channel represented by signal "x" and neighboring channel represented by signal "y". Each channel in both focal and non focal consist of 10240 samples. The focal "x" series, focal "y" series, non focal "x" series and non focal "y" series were represented by Fx, Fy, Nx, and Ny respectively. The sample plot of focal and non focal EEG signal is shown in Fig. 2. From the available four different EEG time series, five classification tasks were formed as follows.

1. Focal "x" series vs Non focal "x" series: FxNx
2. Focal "y" series vs Non focal "y" series: FyNy
3. Focal "x" series vs Non focal "y" series: FxNy
4. Focal "y" series vs Non focal "x" series: FyNx
5. Focal "x" and "y" series vs Non focal "x" and "y" series: FxyNxy

Feature extraction

Feature extraction is one of the important steps in classification problems. Totally, 26 features were considered from different domains such as time, frequency, information theory and statistically based on their previous best performance [6, 23, 30] in seizures detection. Since the database is larger, 10 seconds segmentation was introduced. Table 1 list the 26 EEG features estimated for each segment for each of EEG signal series. The features like mean, variance and standard deviation examined in the first and second derivative of EEG signal [23].

Feature selection

Among 26 features, significant features were selected using Wilcoxon rank sum test by setting p-value and z-score at 95% significance level. A p-value indicates the condition to reject the null hypothesis in favor of the alternative hypothesis. A p-value close to 0 or $< 0.05\%$, shows features can be classified using classifier with resulting good classification result. A z-score is a measure of how many standard deviations below or above the population mean a raw score is (http://www.statisticshowto.com/probability-and-statistics/z-score/). A z-score for confidence interval of 95% is $-1.96 > z > 1.96$. Positive z-score, which indicates the observed value is above the

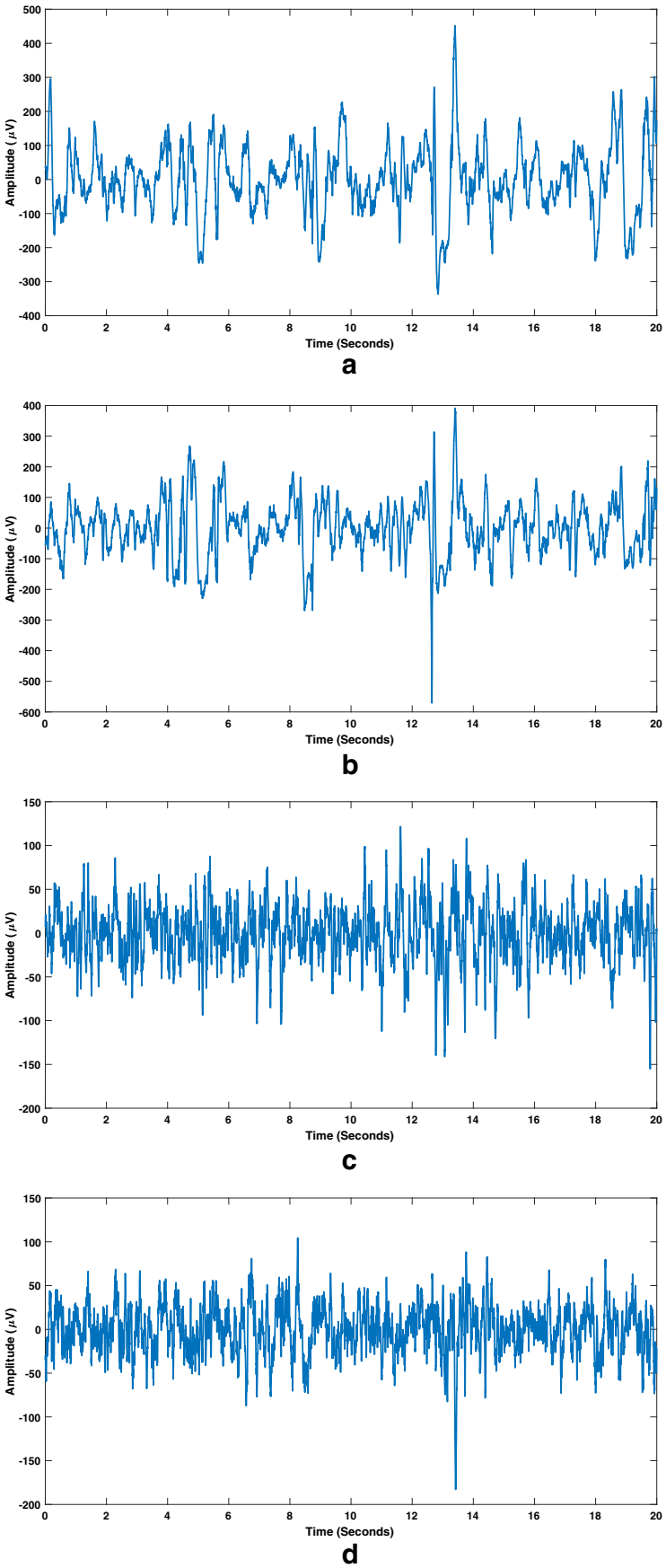

Fig. 2 Plot showing for focal and non focal EEG signals **a** Fx EEG signal, **b** Fy EEG signal, **c** Nx EEG signal, and **d** Ny EEG signal

Table 1 List of EEG features extracted for each segment

EEG feature name	Description
Log energy entropy (Hlog) [30]	It is a non-normalized energy based entropy.
Median frequency (MDF) [1]	EEG power spectrum is divided into two regions with equal amplitude or half of the total power
Mean frequency (MNF) [1]	Sum of product of the EEG power spectrum and the frequency divided by the total sum of the power spectrum
Katz fractal dimension (KFD)[1]	It is an index for characterizing fractal patterns by quantifying their complexity as a ratio of the change in detail to the change in scale
Lower quartile 1 (Q1) [1]	25% of the EEG signal
Upper quartile 3 (Q3) [1]	75% of the EEG signal
Inter quartile range (IQR) [1]	Difference between Q3 and Q1
Semi inter quartile deviation (SID) [1]	It is a measure of spread. It is a half difference between the Q3 and Q1
Skewness (Sk) [5]	Skewness is a measure of degree of symmetry of EEG signal
Kurtosis (Kr) [5]	Kurtosis is a measure of tailedness of the probability distribution of EEG signal
Root mean square (RMS) [17]	Root mean square of the EEG signal
Band power (PB) [17]	Average power in the EEG signal (0-fs/2)
Zero crossing (ZC) [17]	Number of times EEG signal changes its sign
Complexity (Comp) [17]	Hjorth parameter
Mobility (Mob)	Hjorth parameter
Curve length (CL) [2]	Length along a EEG signal
Teager energy (TE) [2]	Non linear energy
Variance (Var) [2]	Variance of the EEG signal
Standard deviation (Std) [2]	Standard deviation of the EEG signal
Mean (Mean) [2]	Mean of the EEG signal
1st derivative variance (Var_1) [2]	Variance of the first derivative of the EEG signal
1st derivative standard deviation (Std_1) [2]	Standard deviation of the first derivative of the EEG signal
1st derivative mean (Mean_1) [2]	Mean of the first derivative of the EEG signal
2nd order derivative variance (Var_2) [2]	Variance of the second derivative of the EEG signal
2nd order derivative standard deviation (Std_2) [2]	Standard deviation of the first derivative of the EEG signal
2nd order derivative mean (Mean_2) [2]	Mean of the first derivative of the EEG signal

mean of all values while negative z-score indicates the observed value is below the mean of all values. A p-value and z-score analysis for each feature is listed in Table 2. The features having $p > 0.05$ and $-1.96 > z < 1.96$ were excluded from the feature set prior to a further step.

Outliers removal

An outlier is a data point that is distinctly separate from the rest of the data. An outlier may be due to variability in the measurement or it may indicate the experimental error. The outliers detection and removal approach reduce the variance of the training data [22]. The studies [21, 22, 37], have reported the improved classification results by removing outliers for EEG signal classification without losing diagnostic information. Turkey's range test was applied to feature set to remove outliers. Initially, observations/data from each feature was sorted in ascending order. The Q1 is the middle value in the first half of the rank-ordered data set and it is also called as $25^{\text{th}}\%$ percentile. The Q3 is the middle value in the second half of the rank-ordered data set and it is also called as $75^{\text{th}}\%$ percentile. The IQR is the difference between Q3 and Q1. According to Turkey's range test [22], an outlier is an observation which is outside the range of $[Q1 - k * IQR, Q3 + k * IQR]$, where k is non-negative integer. Generally, k is set to 1.5 for confidence interval of 95% but it may lead to loss of diagnostic information. In present study, k was set to 3, which indicates far

Table 2 *p*-value and z-score analysis using Wilcoxon rank sum test

	FxNx		FyNy		FxNy		FyNx		FxyNxy	
Feature name	*p*-value	z-score	*p*-value	z-score	*p*-value	z-score	*p*-value	z-score	*p*-value	z-score
Hlog	5.36e-10	6.2082	8.73e-53	15.25	1.14e-56	15.8634	1.13e-51	15.1237	5.15e-106	21.8688
MDF	1.56e-29	− 11.284	6.79e-137	− 24.9037	6.63e-135	− 24.7193	7.34e-130	− 24.2457	3.66e-263	− 34.6557
MNF	3.92e-41	− 13.432	2.84e-199	− 30.1177	8.52e-204	− 30.4612	2.12e-196	− 29.8976	0	− 42.7748
KFD	9.97e-10	− 6.1148	2.73e-29	− 11.2355	8.61e-27	− 10.7155	1.03e-30	− 11.5213	1.19e-55	− 15.7151
Q1	3.72e-09	− 5.89	3.32e-58	− 16.0838	1.67e-62	− 16.6857	1.11e-57	− 16.0086	3.19e-117	− 23.0165
Q3	1.70e-09	5.64	4.50e-39	13.0763	1.17e-42	13.6899	1.02e-36	12.6574	1.84e-77	18.63
IQR	9.01e-11	6.48	2.48e-48	14.6084	3.72e-52	15.1967	7.93e-47	14.3705	1.63e-96	20.8467
SID	9.01e-11	6.48	2.48e-48	14.6084	3.71e-52	15.1968	7.91e-47	14.3706	1.64e-96	20.8466
Sk	0.0237	2.2628	6.89e-16	8.0725	4.55e-10	6.2337	2.09e-15	7.9361	1.20e-24	10.2488
Kr	1.16e-40	13.3518	0.0046	2.8363	6.50e-04	3.4099	0.0108	2.5495	5.63e-05	4.0279
RMS	4.59e-11	6.5838	2.19e-40	13.304	1.31e-43	13.8481	1.54e-38	12.9823	2.96e-79	18.8496
PB	2.94e-11	6.6497	4.66e-40	13.2476	7.76e-43	13.7195	1.22e-36	12.6433	4.38e-76	18.4594
ZC	1.67e-42	− 13.66	7.63e-274	− 35.3577	3.01e-279	− 35.7077	8.11e-274	− 35.356	0	− 50.2911
Comp	9.99e-04	3.2995	1.06e-08	5.7213	3.67e-06	4.6295	1.12e-09	6.0907	1.09e-14	7.7287
Mob	4.20e-34	− 12.17	5.77e-126	− 23.8735	1.78e-132	− 24.4924	1.02e-125	− 23.8495	1.81e-257	− 34.2756
CL	9.36e-09	5.742	2.87e-48	14.5986	2.80e-52	15.2153	6.44e-48	14.5433	5.80e-97	20.8962
TE	0.005	− 2.8039	0.0054	− 2.7813	0.0187	− 2.3521	0.0119	− 2.5148	2.16e-04	− 3.6995
Var	4.89e-11	6.57	2.11e-40	13.3068	1.27e-43	13.85	1.49e-38	12.9847	2.79e-79	18.8527
Std	2.53e-11	6.6714	8.29e-44	13.8807	1.46e-47	14.4871	7.07e-42	13.5583	1.58e-86	19.7157
Mean	0.3973	0.8464	0.3335	− 0.967	0.504	− 0.6682	0.3855	− 0.8678	0.2741	− 1.0937
Var_1	0.0198	− 2.3296	0.0118	− 2.5193	0.0429	− 2.025	0.0245	− 2.2496	0.0011	− 3.254
Std_1	0.0139	− 2.4592	0.0105	− 2.5598	0.053	− 1.9347	0.0144	− 2.4478	0.0012	− 3.2367
Mean_1	0.4657	0.7295	0.2396	− 1.176	0.3029	− 1.0303	0.7206	− 0.3577	0.3739	− 0.8892
Var_2	0.0585	− 1.8918	0.1572	− 1.4145	0.2922	− 1.0534	0.2541	− 1.1405	0.1261	− 1.5297
Std_2	0.0581	− 1.8952	0.3636	− 0.9085	0.6238	− 0.4904	0.4396	− 0.7729	0.393	− 0.8542
Mean_2	0.0887	1.7025	0.793	0.2624	0.9141	0.1079	0.4148	− 0.8154	0.9051	− 0.1192

out data to reduce any loss of information in EEG features. In this study, outliers were tracked and removed prior to a further step. Outliers removed features were concatenated and standardized to transform it to have zero mean and unit variance. Standardization enhance data quality, better data integration, data consistency and clear.

Classification

The efficiency of the pattern classification relies on the selection of appropriate classifier. Apart from sensitivity and specificity, attributes such as speed of learning w.r.t number of features, tolerance to highly independent variable, speed of classification contributes significantly towards the selection of classifier. In the recent attempts on focal and non-focal classification problems, SVM classification found to be showing promising results, that motivated us to choose for classification.

SVM classifier

As the design of SVM is oriented towards functional margin determination, it is found to be a potential candidate for pattern classification problems due to its regularization parameter to avoid over fitting, choice of appropriate kernel, and convex optimization. It has been shown in the literature that the SVM outperforms other linear and nonlinear classifiers in terms of classification accuracy, speed of classification, and tolerance to irrelevant attributes. SVM deliver a unique solution, since the optimality problem is convex. This is an advantage compared to neural networks, which have multiple solutions associated with local minima and for this reason may not be robust over different samples.

SVM is a supervised binary classifier, in this kernel function is used to find the best hyperplane, which separates the training samples of binary class [9, 19]. The best hyperplane for an SVM means the one with the largest margin

between the two classes. The support vectors are the data points that are closest to the separating hyperplane [14, 32]. In this work, quadratic kernel function was found to be the best hyperplane. The quadratic function is also called polynomial kernel (K) which is defined as (http://www.cs.tufts.edu/roni/Teaching/CLT/LN/lecture18.pdf),

$$K(X, Y) = (\gamma \cdot X^T Y + r)^d, \gamma > 0 \quad (1)$$

Here X and Y are vectors in the input space and r, d, and γ are kernel parameters.

Optimization of SVM classifier

Optimization is a process of discovering a point that reduces a real-valued function referred as objective function [20]. Bayesian optimization algorithm was proven to be the best function to minimize the cross-validation loss of a classifier [7, 16]. Bayesian optimization uses Gaussian model of objective function to train model. To estimate minimum objective, 30 objective function evaluations were set with 10-fold cross validation. Figure 3 shows the SVM classifier optimization procedure.

Classifier performance

The classifier performance was evaluated using following measurements [15, 26]. The terminologies derived from confusion matrix are: true positive (TP), false negative (FN), true negative (TN) and false positive (FP). Using four parameters various statistical measures of the performance of a binary classification test was derived.

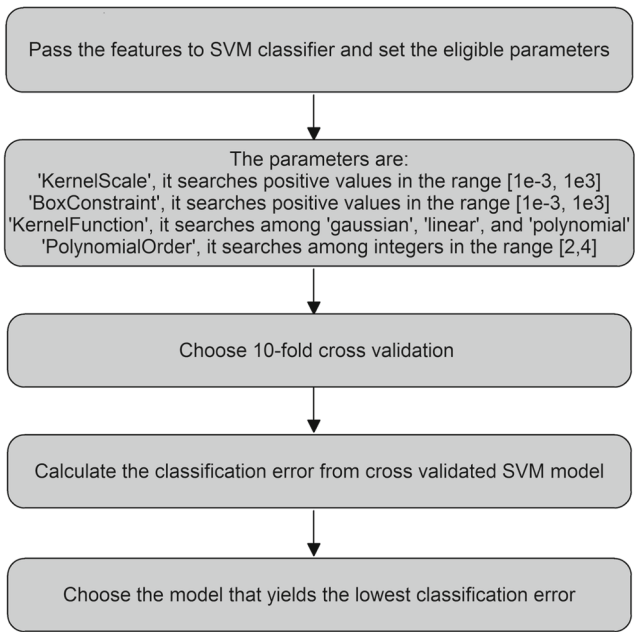

Fig. 3 Support vector machine classifier optimization procedure

$$Sensitivity(SE) = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity(SP) = \frac{TN}{TN + FP} \quad (3)$$

$$Classification\ accuracy(CA) = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

$$Positive\ predictive\ value(PPV) = \frac{TP}{TP + FP} \quad (5)$$

$$Negative\ predictive\ value(NPV) = \frac{TN}{TN + FN} \quad (6)$$

$$False\ positive\ rate(FPR) = 1 - SP \quad (7)$$

$$False\ negative\ rate(FNR) = 1 - SE \quad (8)$$

$$False\ discovery\ rate(FDR) = 1 - PPV \quad (9)$$

$$F1\ score(F1) = \frac{2TP}{2TP + FP + FN} \quad (10)$$

$$Matthews\ correlation\ coefficient(MCC) = \frac{TP*TN - FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (11)$$

$$Informedness(I) = SE + SP - 1 \quad (12)$$

$$Markedness(M) = PPV + NPV - 1 \quad (13)$$

$$Positive\ likelihood\ ratio(LR+) = \frac{SE}{FPR} \quad (14)$$

$$Negative\ likelihood\ ratio(NR-) = \frac{FNR}{SP} \quad (15)$$

$$Diagnostic\ odds\ ratio(DOR) = \frac{LR+}{NR-} \quad (16)$$

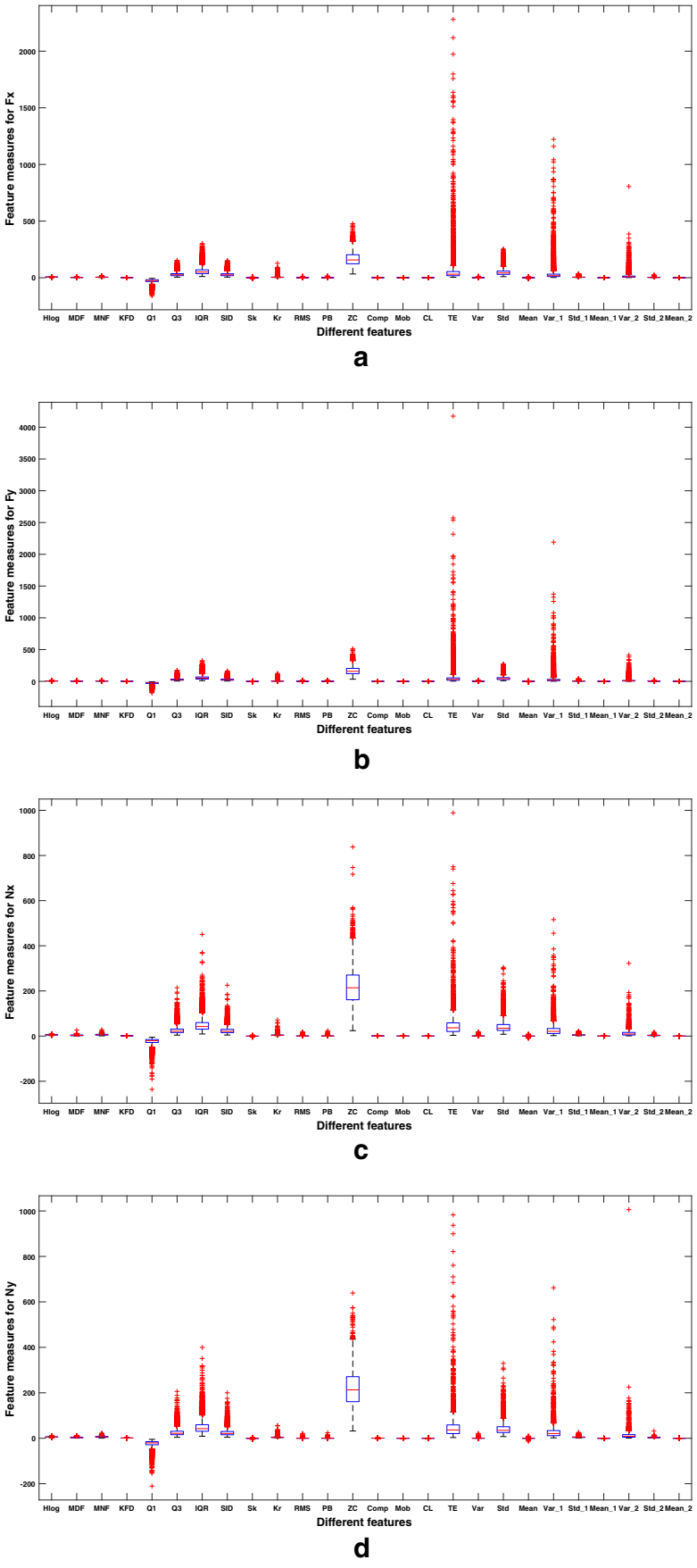

Fig. 4 Boxplot showing for all 26 features for EEG signal series **a** Fx, **b** Fy, **c** Nx, and **d** Ny

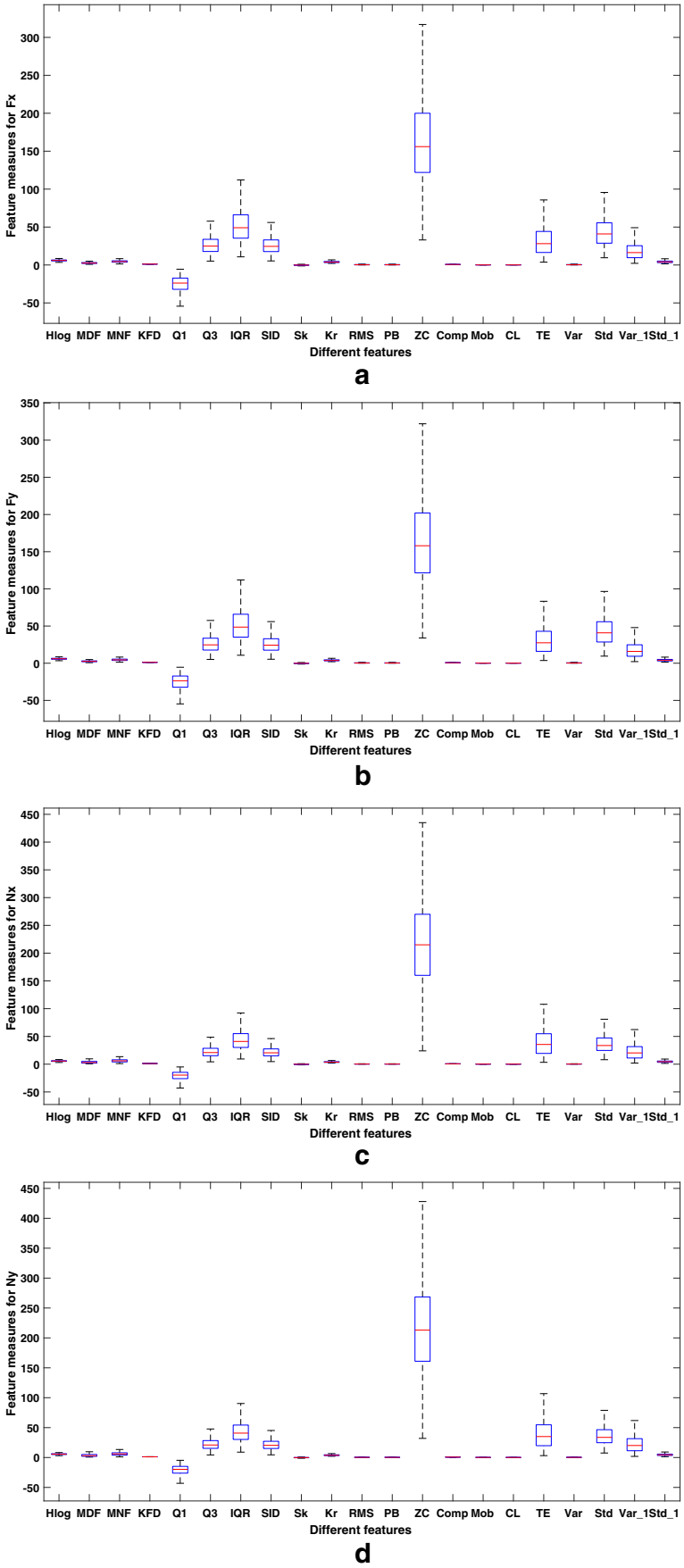

Fig. 5 Boxplot showing for 21 features for EEG signal series **a** Fx, **b** Fy, **c** Nx, and **d** Ny

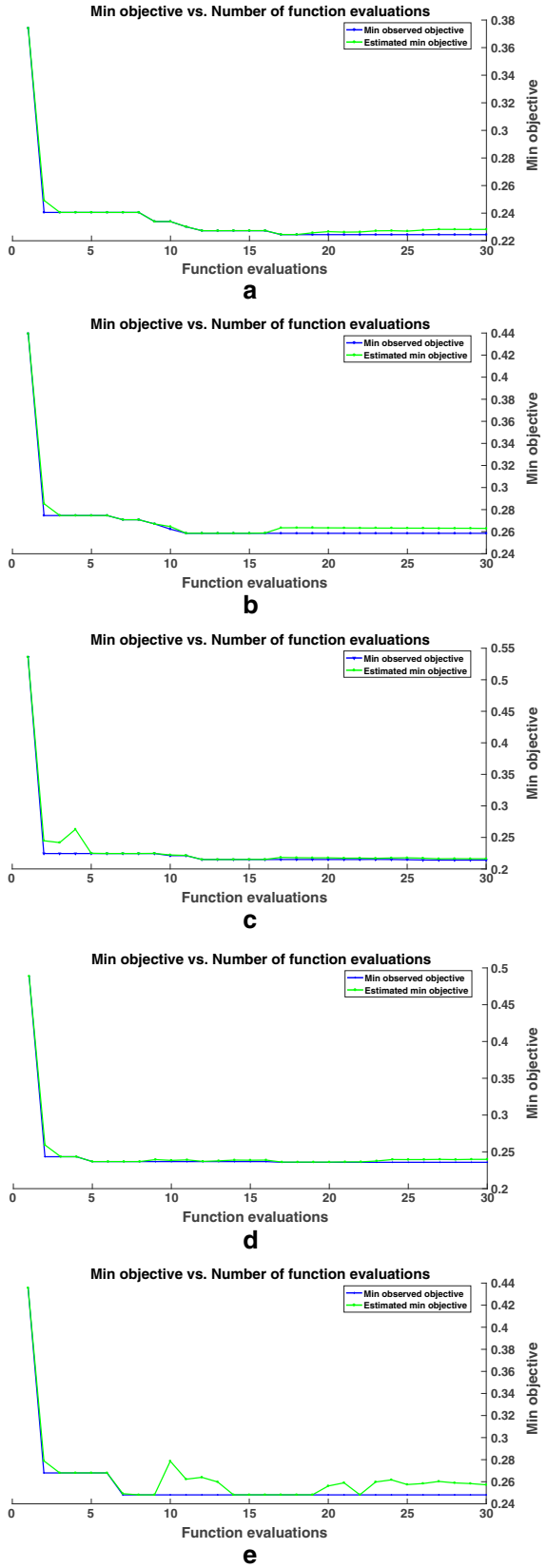

Fig. 6 Plot showing minimum objectives vs. number of function evaluations for different classification task **a** FxNx, **b** FyNy, **c** FxNy, **d** FyNx, and **e** FxyNxy

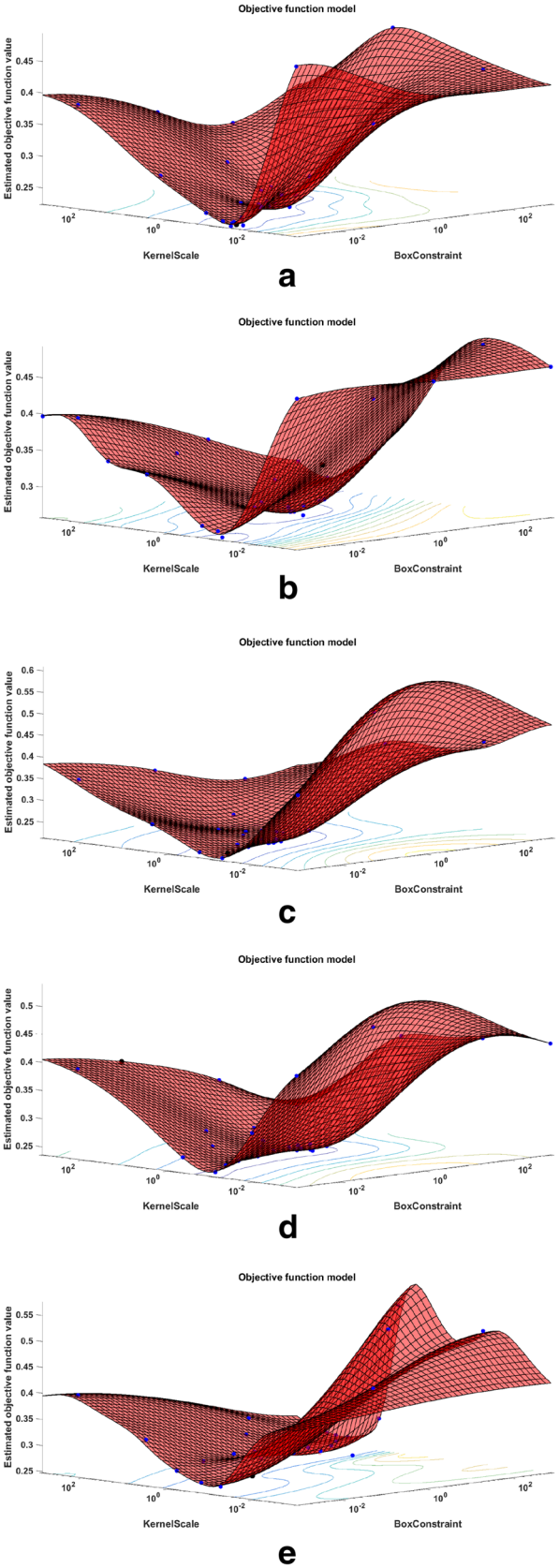

Fig. 7 Plot showing objective function model for classification task **a** FxNx, **b** FyNy, **c** FxNy, **d** FyNx, and **e** FxyNxy

Table 3 SVM optimization results

	FxNx	FyNy	FxNy	FyNx	FxyNxy
Best observed feasible point					
Box Constraint	0.0025	173.07	1.8701	3.6315	0.2666
Kernel Scale	0.0467	8.4416	1.3343	1.5140	0.8737
Observed objective function value	0.2244	0.2481	0.2585	0.2358	0.2140
Estimated objective function value	0.2282	0.2572	0.2629	0.2398	0.2161
Function evaluation time	100.30	108.26	100.31	105.10	240.93
Kernel	Quadratic	Quadratic	Quadratic	Quadratic	Quadratic
Best estimated feasible point (according to models)					
BoxConstraint	0.0015	24.107	1.3060	3.3150	0.2666
Estimated function evaluation value	86.607	53.844	81.778	104.18	240.59

The performance parameters such as SE, SP, CA, PPV, NPV, F1, MCC, I, M, LR+, and DOR expected to be higher values while remaining parameters considered to fall under lower bound to ensure the suitability of the proposed approach for clinical routine.

Results and discussion

Simulation results

The EEG signals were segmented into 10 seconds segment to compute each feature. All 26 features are shown in Fig. 4 using boxplot. Outliers can be observed in each of the features and it degrade classifier performance. From the Fig. 4, the features such as TE and Var_1 has more outliers for focal EEG signal, whereas Q1, Q3, IQR, SID, ZC, TE, Std, Var_1, and Var_2 has more outliers for non focal EEG signal.

Wilcoxon rank sum test was performed on each of the classification task to identify the significant features. A p-value and z-score analysis for different classification task is shown in Table 2. The confidence interval was set to 95% indicating that the features with $p < 0.05$ and $-1.96 > z > +1.96$ can be considered as significant features for classification and features with $p>0.05$ and $-1.96 < z < +1.96$ are highlighted. The above-mentioned condition indicates that there is enough evidence to reject the null hypothesis. The features such Mean, Mean_1, Var_2, Std_2, and Mean_2 fail to obtain typical p and z-score. These five features were excluded from the feature set prior to further analysis based on p-value and z-score.

The studies [11, 33–35], have classified the features in the presence of the outliers. Statistically, considering outliers for analysis and classification is not the intelligent way. After Wilcoxon rank sum test, 21 features were selected. Outliers were removed using Turkey's range test for all 21 features. The features without outliers from every EEG times series are shown in Fig. 5.

The outliers removed feature set was standardized prior training the SVM classifier. An optimized SVM model was arrived by using objective function model and minimum objectives vs. number of function evaluations. Figure 6 shows graph of minimum objectives vs. number of function evaluations for the different classification task. Best objective model function for different classification task is depicted in Fig. 7. Optimization results of SVM classifiers are listed in Table 3. It can be clearly inferred from Table 3 that the best observed feasible point matches with best estimated feasible point and less classification loss was attained for FxyNxy.

The performance of optimized SVM classifier using multi-features is reported in Table 4. Results have been shown individually for each classification task and best results were highlighted. The highest SE, SP, and CA of 0.9456, 0.8975, and 0.9215 was achieved respectively for classification task FxyNxy. The lowest SE of 0.8582, SP

Table 4 Performance of SVM classifier using multi-features

	SE	SP	CA	PPV	NPV	FPR	FNR	FDR	F1	MCC	I	M	LR+	LR-	DOR
FxNx	0.9034	0.8523	0.8778	0.8595	0.8982	0.1477	0.0966	0.1405	0.8809	0.7567	0.7557	0.7577	6.1154	0.1133	53.9527
FyNy	0.8712	0.7595	0.8153	0.7836	0.8550	0.2405	0.188	0.2164	0.8251	0.6347	0.6307	0.6387	3.6221	0.1696	21.3505
FxNy	0.8582	0.7633	0.8106	0.7837	0.8431	0.2367	0.1422	0.2163	0.8192	0.6214	0.6212	0.6268	3.6244	0.1861	19.4730
FyNx	0.8754	0.8371	0.8561	0.8431	0.8701	0.1629	0.1251	0.1569	0.8587	0.7126	0.7121	0.7131	5.3721	0.1493	35.9767
FxyNxy	0.9456	0.8974	0.9215	0.8921	0.9896	0.1296	0.0644	0.1279	0.8974	0.8289	0.8362	0.8117	6.8532	0.0757	87.5786

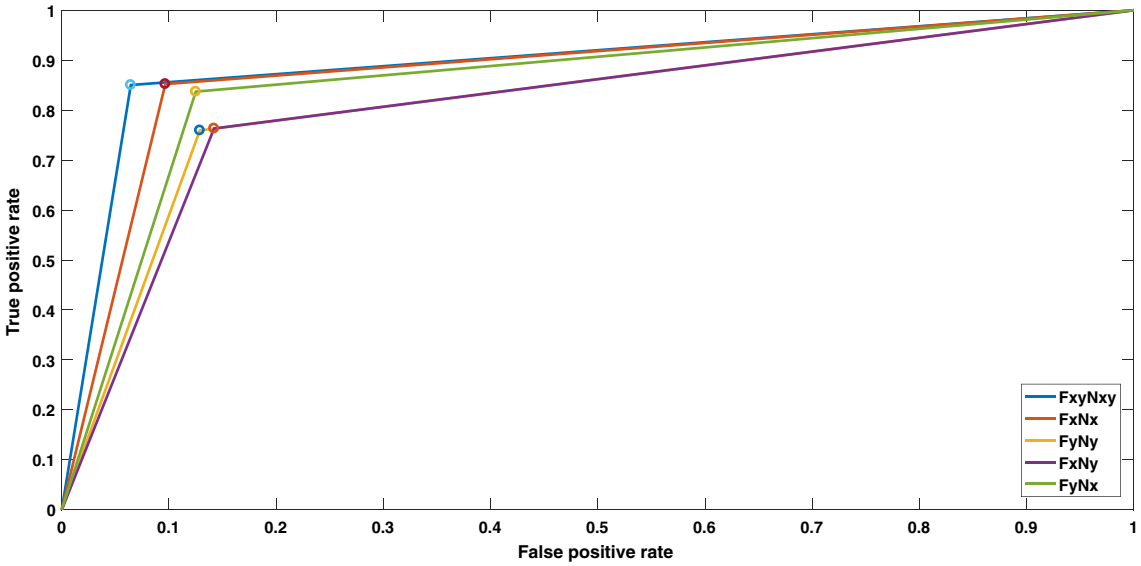

Fig. 8 Receiver operating characteristic curve for focal and non focal EEG signal

of 0.7595, and CA of 0.8153 was achieved for classification task FxNy, FyNy, and FyNy respectively. The highest PPV, NPV, F1, MCC, I, M, LR+, and DOR was attained for classification task FxyNxy and lowest FPR, FNR, FDR, and LR- was attained for FxyNxy. It can be concluded from the simulation results that the classification task FxyNxy has achieved best score among five tasks. The receiver operating characteristic (ROC) curve for different classification tasks is shown in Fig. 8. The highest area under the curve of 0.9483 was recorded for classification task FxyNxy.

The CA of proposed model with and without outliers is listed in Table 5. It was observed that improved CA was obtained without outliers. Maximum test CA difference of 11.89% was achieved without outliers for classification task FxyNxy. Hence, it is evident that removing outliers play an important role in improving classifier performance.

Comparative study

In recent times, some studies has been conducted using the same database for the classification of focal and non focal EEG signal. The CA results of the proposed study compared with other studies using the same database is reported in Table 6. It is clearly evident from the Table 6, that the proposed method obtained highest CA without the usage of any signal decomposition technique. The highest CA was attained by selecting significant features using p-value and z-score. The removal of outliers further ensured a significant improvement in terms of CA (11.89%). The diagnostic quality of the EEG recordings were validated again for clinical decision and the same was confirmed by the specialist in the local hospital. It can be concluded that the proposed combination of multi-features derived from different domain was found to be reliable to discriminate focal and non focal EEG other than a single feature.

In [33], focal and non focal EEG signals were decomposed up to six levels using DWT to compute different entropies in approximate and detail coefficients of sub-band signals. Student's t-test and receiver operating characteristic-based ranking methods achieved average CA of 84%, sensitivity of 84% which is quite lower than our method. In study [11], EMD followed by DWT was applied

Table 5 Comparison of classification accuracy with and without outliers

	FxNx	FyNy	FxNy	FyNx	FxyNxy
With outliers	0.7941	0.7701	0.7635	0.7824	0.8026
Without outliers	0.8778	0.8153	0.8106	0.8561	0.9215
Improvement (%)	8.37	4.52	4.68	7.37	11.89

The difference in the classification accuracy with and without outliers is highlighted

Table 6 Comparison of studies performed on the Bern-Barcelona database for focal and non-focal classification

Authors	No. signals used for study	Signal decomposition method	CA(%)
Chen et al. [8]	750	DWT	83.07
Zhu et al. [38]	50	No	84
Sharma et al. [34]	50	DWT	84
Sharma et al. [33]	50	EMD	85
Sharma et al. [35]	50	EMD	87
Chua et al. [10]	50	EMD	88.78
Das et al. [11]	50	EMD-DWT	89.04
Proposed work	3750	No	92.15

The best results obtained from proposed study are highlighted

on EEG signal to compute log energy entropy. They did not remove outliers from feature samples. The obtained features were classified using k-NN classifier which gave CA of 89.4%. The procedure was quite a complex due to the combination of EMD and DWT. Sharma et al. [34] extracted sample entropies and variances of the intrinsic mode functions obtained by EMD of EEG signals. Using LS-SVM classifier, CA of 85% was attained using two features. In another study [8], EEG signals were decomposed using various wavelet functions such as bior5.5, coif3, db2, dmey, haar, rbio6.8, and sym6 to extract 9 different statistical features. The highest CA of 83% was achieved using SVM with radial basis function as a kernel.

Studies reported in Table 6 have adopted signal decomposition step either as DWT or EMD and both in some studies. Further, outliers were not removed from the features prior to classification. Previous studies had failed to reach the CA of above 90% using a single feature. However, proposed study has succeeded to achieve CA of 92.15% using multi-features without using any signal decomposition technique. One can observe from the study reported, that the optimized SVM technique outperforms the work reported earlier in terms of classification accuracy.

Conclusion

This paper presents multi-features based on localization of focal seizures in pharmacoresistant patients which helps to identify epileptogenic zones prior to surgery. The focal and non focal EEG was collected from Bern Barcelona publicly available database. Initially, 26 features were extracted from 10 seconds segment from focal and non focal, "x" and "y" EEG series. Totally five different classification tasks were formed. Wilcoxon rank sum test was performed to identify significant features as a result 5 features were found to be insignificant for classification. A hypothesis was made that removing outliers would enhance accuracy, hence outliers were removed from rest 21 features using Turkey's range test. Bayesian optimization technique was adapted to ensure the minimal cross-validation loss in SVM classifier. Only 21 significant features were used to classify using SVM classifier using 10-fold cross validation. The proposed approach using multi-features obtained best of 94.56%, 89.74%, and 92.15%, sensitivity, specificity, and accuracy respectively, which is higher than that of the state-of-the-art approaches. The improvement in accuracy of 11.89% was achieved by removing the outliers. The performance of the multi-features with the optimized SVM classifier was evaluated using various classifier parameters and the results confirms the suitability of the proposed scheme for real-time focal seizures detection. The proposed method differs from other methods in literature in terms of multi-features, outliers removal, and Bayesian optimized SVM classifier. It can be concluded that the proposed approach can be applied for identification of focal EEG signals to localize epileptogenic zones.

Acknowledgements Authors would like to thank Dr. R.G. Andrzejak, for providing permission to use EEG database for research work. The authors would also like to thank the anonymous reviewers for their helpful comments and suggestions that greatly improved the quality and clarity of the manuscript.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

Ethical approval The proposed study makes use of open source database where appropriate ethical clearance has been taken.

References

1. Aarabi, A., Wallois, F., and Grebe, R., Automated neonatal seizures detection: A multistage classification system through feature selection based on relevance and redundancy analysis. *Clin. Neurophysiol.* 117:328–340, 2006.
2. Aarabi, A., Grebe, R., and Wallois, F., A multistage knowledge-based system for EEG seizures detection in newborn infants. *Clin. Neurophysiol.* 118(12):2781–2797, 2007.
3. Acharya, U. R., Molinari, F., Vinitha, S. S., and Chattopadhyay, S., Automatic diagnosis of epileptic EEG using entropies. *Biomed. Signal Process. Control* 7(4):401–408, 2012.
4. Andrzejak, R. G., Schindler, K., and Rummel, C., Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients. *Phys. Rev. E.* 2012(86):046206:1–17, 2012.
5. Bedeeuzzaman, M., Farooq, O., and Khan, Y. U., Automatic seizures detection using higher order moments. In: *Proceedings of the International Conference on Recent Trends in Information, Telecommunication and Computing*, pp. 159–163, 2010.
6. Bogaarts, J. G., Gommer, E. D., Hilkman, D. M. W., van Kranen-Mastenbroek, V. H. J. M., and Reulen, J. P. H., Optimal training dataset composition for SVMbased, ageindependent, automated epileptic seizures detection. *Med. Biol. Eng. Comput.* 54:1285–1293, 2016.
7. Bull, A. D.: Convergence rates of efficient global optimization algorithms. 2011. arXiv:1101.3501v3.
8. Chen, D., Wan, S., and Bao, F. S., Epileptic focus localization using EEG based on discrete wavelet transform through full-level decomposition. In: *IEEE International Workshop on Machine Learning for Signal Processing. Sept. 17–20*, 2015.
9. Christianini, N., and Shawe-Taylor, J. C., *An introduction to support vector machines and other kernel-based learning methods*. Cambridge: Cambridge University Press, 2000.
10. Chua, K. C., Chandran, V., Acharya, U. R., and Lim, C. M., Automatic identification of epilepsy by HOS and power spectrum parameters using EEG signals: A comparative study. *Proc. IEEE Eng. Med. Biol. Soc. Conf.* 2008:3824–3827, 2008.
11. Das, A. B., and Bhuiyan, M. I. H., Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain. *Biomed. Signal Process. Control* 29:11–21, 2016.

12. Delaneya, R. C., Alexander, J. R., Mattson, R. H., and Novelly, R. A., Memory function in focal epilepsy: A comparison of non-surgical, unilateral temporal lobe and frontal lobe samples. *Cortex* 16(1):103–117, 1980.
13. Eitricha, T., and Langb, B., Efficient optimization of support vector machine learning parameters for unbalanced datasets. *J. Comput. Appl. Math.* 196(2):425–436, 2006.
14. Fan, R. E., Chen, P. H., and Lin, C. J., Working set selection using second order information for training support vector machines. *J. Mach. Learn. Res.* 6:1889–1918, 2005.
15. Fawcett, T., An introduction to ROC analysis. *Pattern Recogn. Lett.* 27(8):861–874, 2006.
16. Gelbart, M., Snoek, J., and Adams, R. P.: Bayesian optimization with unknown constraints. 2014. arXiv:1403.5607.
17. Greene, B. R., Faul, S., Marnane, W. P., Lightbody, G., Korotchikova, I., and Boylan, G. B., A comparison of quantitative EEG features for neonatal seizures detection. *Clin. Neurophysiol.* 119(6):1248–1261, 2008.
18. Guler, I., and Ubeyli, E. D., Multiclass support vector machines for EEG-signals classification. *IEEE Trans. Inf. Technol. Biomed.* 2(2):117–126, 2007.
19. Kecman, V. *Learning and soft computing*. Cambridge: MIT Press, 2001.
20. Kecman, V., Huang, T. M., and Vogt, M., Iterative single data algorithm for training kernel machines from huge data sets: theory and performance. In: *Support Vector Machines: Theory and Applications*, pp. 255–274: Springer, 2005.
21. Krauledat, M., Dornhege, G., Blankertz, B., and Muller, K. B., Robustifying EEG data analysis by removing outliers. *Chaos and Complexity Theory* 2(2):259–274, 2007.
22. Li, W., Mo, W., and et al, Outlier detection and removal improves accuracy of machine learning approach to multispectral burn diagnostic imaging. *J Biomed Opt.* 20(12):121305, 2015.
23. Menshawy, M. L., Benharref, A., and Serhani, M., An automatic mobile-health based approach for EEG epileptic seizures detection. *Expert Systems with Applications* 42:7157–7174, 2015.
24. Moshe, S. L., Perucca, E., Wiebe, S., and Mathern, G. W.: The International League Against Epilepsy at the threshold of its second century, 2011.
25. Pauri, F., Pierelli, F., Chatrian, G. M., and Erdly, W. W., Long-term EEG-video-audio monitoring: computer detection of focal EEG seizures patterns. *Electroencephalogr. Clin. Neurophysiol.* 82(1):1–9, 1992.
26. Powers, D. M. W., Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *J. Mach. Learn. Technol.* 2(1):37–63, 2011.
27. Pravin, S., Sriraam, N., Benakop, P., and Jinaga, B., Entropies based detection of epileptic seizures with artificial neural network classifiers. *Expert System Applications* 37(4):3284–3291, 2010.
28. Ozbeyaz, A., Gürsoy, M. I., and Oban, R., Regularization and kernel parameters optimization based on PSO algorithm in EEG signals classification with SVM. In: *2011 IEEE 19th Signal Processing and Communications Applications Conference (SIU), Antalya*, pp. 399–402, 2011.
29. Raghu, S., Sriraam, N., and Pradeep Kumar, G., Effect of wavelet packet log energy entropy on electroencephalogram (EEG) signals. *International Journal of Biomedical and Clinical Engineering* 4(1):32–43, 2015.
30. Raghu, S., Sriraam, N., and Pradeep Kumar, G., Classification of epileptic seizures using wavelet packet log energy and norm entropies with recurrent Elman neural network classifier. *Cognitive Neurodynamics Springer* 11(1):51–66, 2016.
31. Raghu, S., and Sriraam, N., Optimal configuration of multilayer perceptron neural network classifier for recognition of intracranial epileptic seizures. *Expert Systems With Applications* 89:205–221, 2017.
32. Scholkopf, B., and Smola, A. *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond, Adaptive Computation and Machine Learning*. Cambridge: The MIT Press, 2002.
33. Sharma, R., and Pachori, A. U. R., Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals. *Entropy* 17:669–691, 2015.
34. Sharma, R., Pachori, R. B., and Acharya, U. R., An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures. *Entropy* 17:5218–5240, 2015.
35. Sharma, R., Pachori, R. B., and Gautam, S., Empirical mode decomposition based classification of focal and non-focal EEG signals. In: *Proceedings of 2014 International Conference on Medical Biometrics, Shenzhen, China*, pp. 135–140, 2014.
36. Srinivasan, V., Eswaran, C., and Sriraam, N., Approximate entropy-based epileptic EEG detection using artificial neural networks. *IEEE Trans. Inf. Technol. Biomed.* 11(3):288–295, 2007.
37. Winkler, I., Haufe, S., and Tangermann, M., Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioural and Brain Functions* 7(30):1–35, 2007.
38. Zhu, G., Li, Y., Wen, P. P., Wang, S., and Xi, M., Epileptogenic focus detection in intracranial EEG based on delay permutation entropy. In: *Proceeding of 2013 International Symposium on Computational Models for Life Science, Sydney, Australia*, Vol. 1559, p. 3136, 2013.