Automated Classification of Focal and Non-focal EEG Signals using Fourier-Bessel Series Expansion

A PROJECT REPORT

Submitted in partial fulfillment for the
Degree of Bachelor of Technology

Discipline of Electrical Engineering

Submitted by
Swastik Gupta (140002036)
Konduri Hari Krishna (140002017)

Project Supervisors – Dr. Ram Bilas Pachori
Dr. M. Tanveer

INDIAN INSTITUTE OF TECHNOLOGY INDORE

December, 2017

Declaration by the Candidate

It is hereby declared that the project entitled "**Automated Classification of Focal and Non-focal EEG Signals using Fourier-Bessel Series Expansion"** submitted in partial fulfillment for the award of the degree of Bachelor of Technology in Electrical Engineering completed under the supervision of **Dr. Ram Bilas Pachori, Associate Professor, Discipline of Electrical Engineering, IIT Indore** and **Dr. M. Tanveer, Assistant Professor and Ramanujan Fellow, Discipline of Mathematics, IIT Indore**, is an authentic work.

Further, this work has not been submitted for the award of any other degree elsewhere.

Konduri Hari Krishna	Swastik Gupta
Roll no - 140002017	Roll no – 140002036

Certificate by Project Supervisors

It is certified that the above statement made by the students is correct to the best of our knowledge.

Dr. Ram Bilas Pachori	Dr. M. Tanveer
Associate Professor	Assistant Professor and Ramanujan Fellow
Discipline of Electrical Engineering	Discipline of Mathematics
Indian Institute of Technology Indore	Indian Institute of Technology Indore

Preface

This project on "Automated Classification of Focal and Non-focal EEG Signals using Fourier-Bessel Series Expansion" is prepared under the supervision of Dr. Ram Bilas Pachori and Dr. M. Tanveer.

An entirely novel and unprecedented method was theorized and implemented for the two class classification problem of focal and non-focal electroencephalogram (EEG) signals. In this classification problem we have involved various steps such as computation of difference of bivariate EEG signals, signal decomposition, features extraction and classification using a binary classifier.

This project comprehensively explains the entire implementation of the method in a step-by-step manner. The difference of bivariate EEG signals increase the discrimination between two class of focal and non-focal EEG signals. Fourier-Bessel series expansion technique is used to decompose these difference EEG signals. In the features extraction process, we have computed 17 different features for classification purpose. Support vector machine (SVM) is used for the classification of focal and non-focal EEG signals along with its least square formulation. The proposed method shows a maximum classification accuracy of 84.25%. The automated detection system can also be implemented for successful clinical purpose for the classification of focal and non-focal EEG signals.

Swastik Gupta (140002036)

Konduri Hari Krishna (140002017)

B.Tech. 4th Year

Discipline of Electrical Engineering

Indian Institute of Technology Indore

Acknowledgements

We would like to thank Dr. Ram Bilas Pachori for his valuable guidance in this project, enlightening us with useful insights throughout its duration.

We would also like to thank Dr. M. Tanveer for his crucial assistance in the field of machine learning. His guidance played an important role in the implementation of binary classifiers.

There are many others who share the effort that has been put in this project, simply because its completion would have never been possible without their help. Having said that, we would also like to express our gratitude to Mr. Vipin Gupta, Research Associate, Signal Analysis Research Lab, who assisted us throughout the project.

Finally, we would like to sincerely thank everyone else who knowingly or unknowingly helped us in the project.

Swastik Gupta (140002036)

Konduri Hari Krishna (140002017)

B.Tech. 4th Year

Discipline of Electrical Engineering

Indian Institute of Technology Indore

Abstract

In this project, we have proposed a new method for automated classification of focal and non-focal electroencephalogram (EEG) signals. The dataset used in our work contains bivariate EEG signals of both focal and non-focal classes. These signals are first employed to compute the difference of bivariate EEG signals. The difference of bivariate EEG signals is computed to increase the discrimination between focal and non-focal classes.

The difference of bivariate signals is decomposed using Fourier-Bessel series expansion. The coefficients from this decomposition process are further segmented into 5 parts. These 5 small segments of the Fourier- Bessel series coefficients are considered for the features extraction process in which 17 different features are computed.

The extracted features from focal and non-focal EEG signals are used for classification. The binary classifiers such as support vector machine (SVM) and least square support vector machine (LS-SVM) along with various kernel functions such as linear, polynomial, and radial basis function (RBF) have been implemented in our work for the comparison of obtained classification accuracies. A 10-fold cross-validation method is used to verify the performance of the classifiers in term of accuracy.

Contents

List of Figures and Tables

List of Abbreviations

EEG	Electroencephalogram
FB	Fourier-Bessel
SVM	Support vector machine
LS-SVM	Least square support vector machine
RBF	Radial basis function
IMF	Intrinsic mode function
EMD	Empirical mode decomposition
EWT	Empirical wavelet transform
KNN	K-nearest neighbor
TQWT	Tunable-Q wavelet transform
FE	Focal EEG
NFE	Non-focal EEG

Chapter 1. Introduction

The human body relies on the electrical signals that are generated in the brain for the entirety of its functions. Abnormal neuronal activity in the brain results in epileptic seizure and occurrence of at least one such seizure may results in epilepsy [1]. Generalized epilepsy is the one where the whole part of the brain is affected and the other one is partial or focal epilepsy which involves localized epileptic discharge [2]. Epilepsy is the most common neurological disorder and more than 50 million patients are suffering from its worldwide [3]. Seizures involved with focal epilepsy cannot be controlled with medications and the only way to cure is to locate the focal epileptogenic area and surgically remove it [4].

The electroencephalogram (EEG) signals measure the electrical activity of the brain. The EEG signals can be recorded in two manners, the first one is scalp and another one is intracranial EEG recordings [5]. Generally, intracranial EEG recording is used for the purpose of locating the focal epileptogenic zone [6]. Various computer-aided detection techniques have been developed to understand the dynamics based on the EEG signals and the techniques employed in these methods involve advanced signal processing methods and they are useful for the localization of the epileptogenic focal area of the brain [7-15]. In the literature, average sample entropies and average variance of instantaneous frequencies features computed from intrinsic mode functions (IMFs) extracted by empirical mode decomposition (EMD) have been fed to least square support vector machine (LS-SVM) classifier with radial basis function (RBF) kernel. A classification accuracy of 85% has been achieved by this method for the focal and non-focal EEG signals [14]. In another work [15], average entropy features like Shannon wavelet, fuzzy, Tsallis wavelet, Renyi wavelet associated with EMD has yielded a classification accuracy of 87% on 50 EEG signals database. In [13], the EMD is being associated with log-energy entropy feature and the usage of K-nearest neighbor (KNN) classifier has resulted in a classification accuracy of 89.4% on 3750 EEG signals database. In another work [7], the empirical wavelet transform (EWT) has been deployed to decompose both focal and non-focal EEG signals into rhythms then the areas obtained from the reconstructed phase space plot of rhythms corresponding to different central tendency measures are used as features with LS-SVM classifier. The obtained classification accuracies are 90% and 82.53% for 50 and 750 focal and non-focal EEG signals, respectively. The decision support system based on Tunable-Q wavelet transform (TQWT) method has been proposed in [9] which resulted in a classification accuracy of 95% for full dataset.

The discrimination between focal and non-focal EEG signals is a difficult and time consuming task by visual inspection. Thus, a signal decomposition method is imperative to effective study of these EEG signals and classifies them. In this work, the main focus is to propose a new methodology for automated classification of focal and non-focal EEG signals using Fourier-Bessel series expansion. The Fourier-Bessel (FB) series coefficients of bivariate EEG signal is further segmented into 5 parts for the features extraction process and 17 distinguishing features are computed. These features are used for the classification of focal and non-focal EEG signals using two classifiers namely SVM and LS-SVM.

The methodology used in our work is entirely unprecedented and has been used for the first time for binary classification of EEG signals. The different steps involved in the project can be understood through the block diagram sequentially.

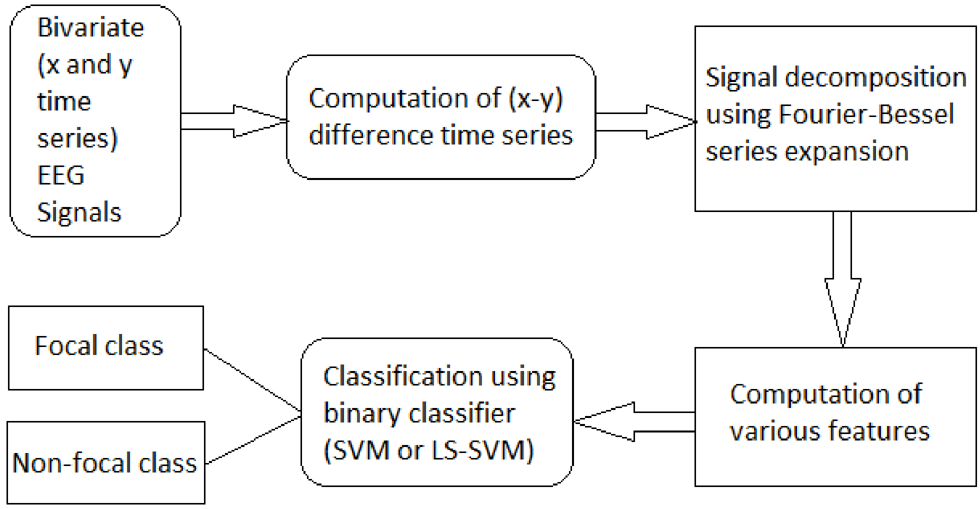

Figure 1.1. Block diagram of the proposed method.

Chapter 2. Methodology

The entire methodology of our work is divided into four major sections. Section 2.1 provides the database description and signal processing that has to be done before performing signal decomposition. Section 2.2 explains the signal decomposition method implemented in our work. The decomposition of bivariate EEG signals is followed by the extraction of features, which is explained in section 2.3. The last but a vital component of our methodology involves implementation of binary classifiers for training and testing of these EEG signals based on the features extracted. The different classifiers used in our work have been explained in sub-sections 2.4.1 and 2.4.2.

2.1. Dataset Description

The dataset used in our work is obtained from (www.dtic.upf.edu/ralph/sc/) which is a publicly available database consisting of two-channel intracranial EEG recordings of five patients suffering from drug-resistant focal epilepsy. This dataset consists of bivariate (x and y) EEG signals. Each bivariate EEG signal has 10,240 samples with a sampling frequency of 512 Hz [16]. Figures 2.1 and 2.2 illustrate the x and y time series of focal EEG signals, Figures 2.3 and 2.4 illustrate the x and y time series of non-focal EEG signals.

In this work, we have used 400 focal and 400 non-focal bivariate EEG signals for the classification purpose. To increase the discrimination between these bivariate EEG signals, an efficient technique is to compute the difference (x-y) of bivariate (x and y) from focal and non-focal EEG signals as suggested by [9, 13]. Figures 2.5 and 2.6 represent the difference (x-y) time series of focal and non-focal EEG signals respectively.

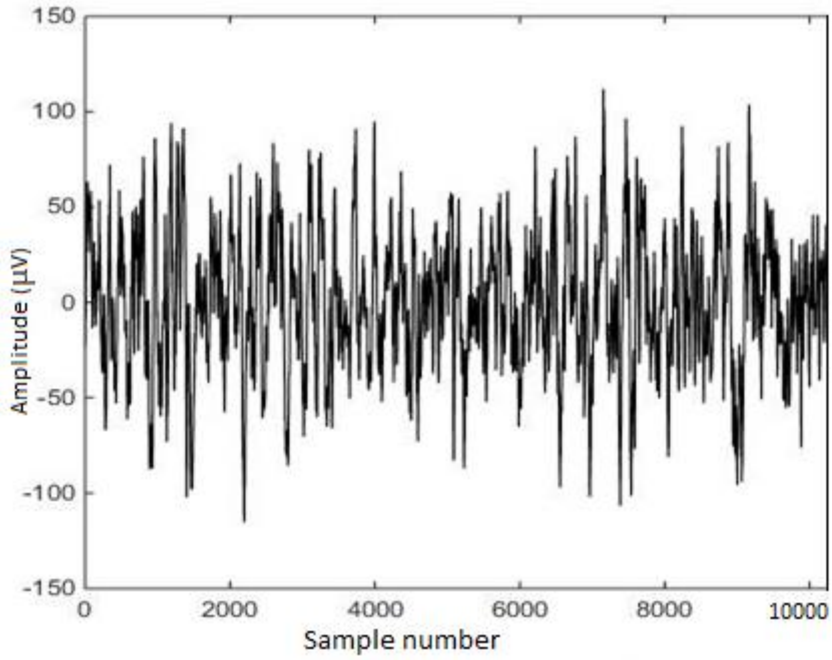

Figure 2.1. x-time series of focal EEG signal.

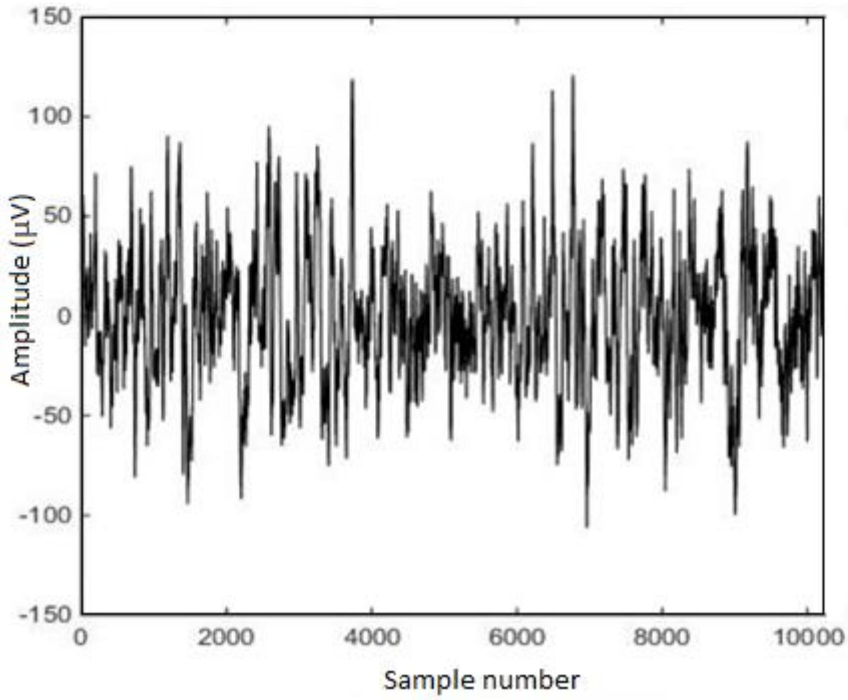

Figure 2.2. y-time series of focal EEG signal.

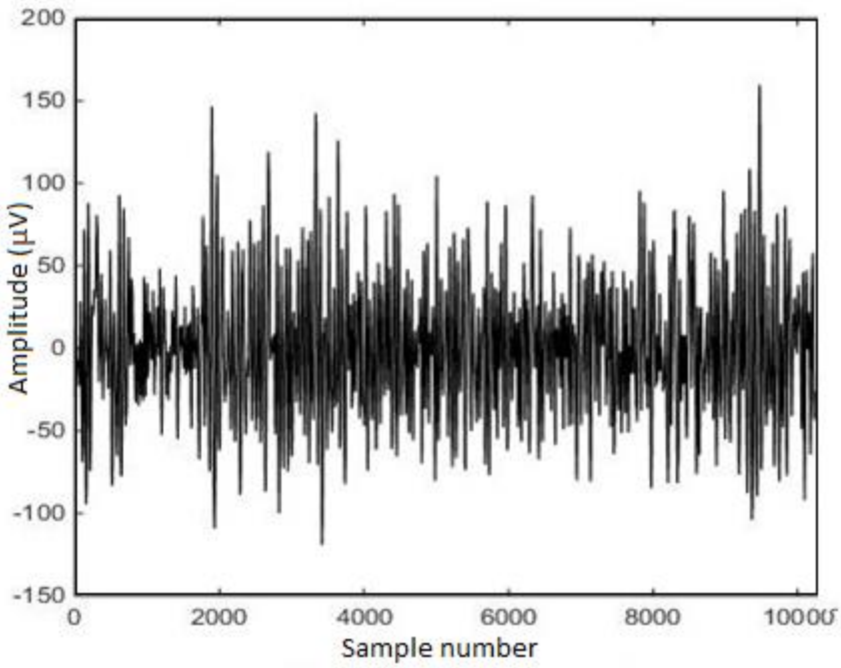

Figure 2.3. x-time series of non-focal EEG signal.

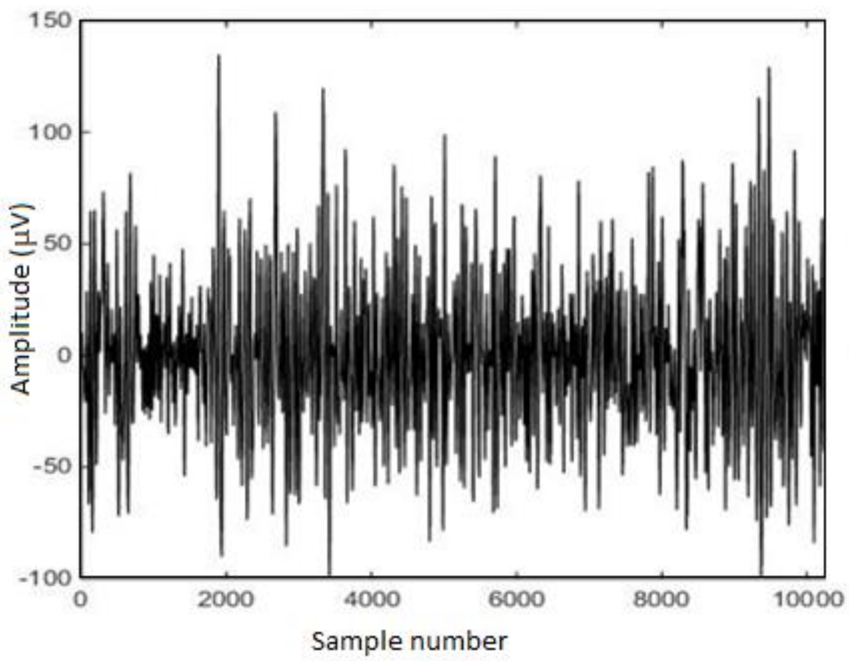

Figure 2.4. y-time series of non-focal EEG signal.

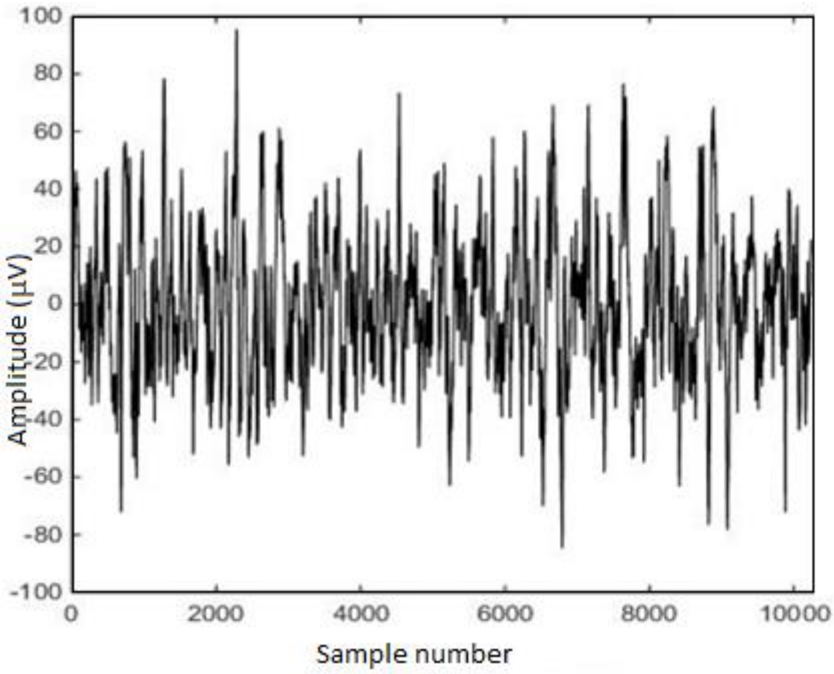

Figure 2.5. Difference (x-y) time series of focal EEG signals

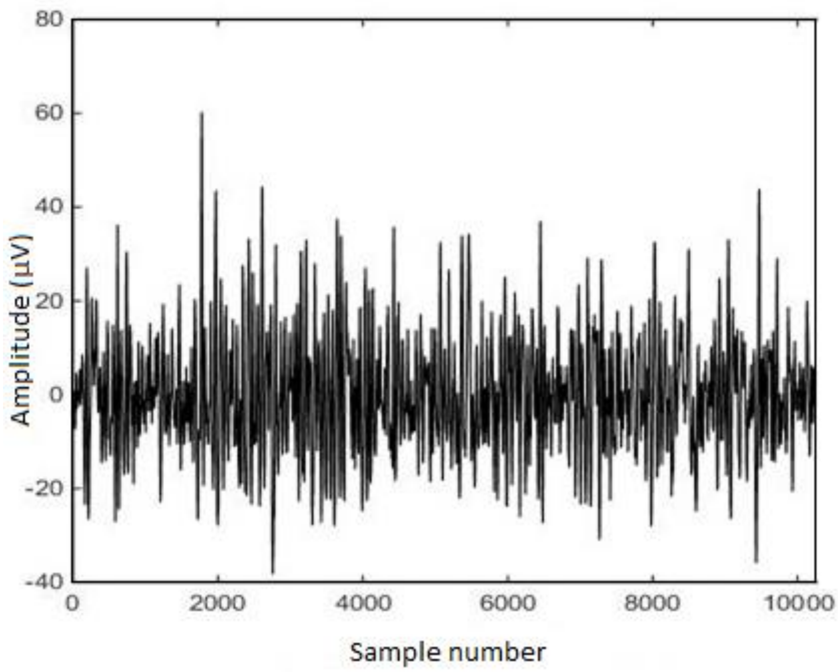

Figure 2.6. Difference (x-y) time series of non-focal EEG signals

2.2. Fourier-Bessel Series Expansion

Decaying and aperiodic Bessel functions form the basis for the Fourier-Bessel series expansion. The zero-order Fourier-Bessel series expansion of any discrete-time signal *x(n)* over an interval [0,N] can be written as [17],

$$x(n) = \sum_{m=1}^{M} C_m \, J_0\left(\frac{\lambda_m n}{N}\right),$$

The coefficients C_m are calculated as,

$$C_m = \frac{2\int_0^N n x(n) J_0\left(\frac{\lambda_m n}{N}\right) dn}{N^2 [J_1(\lambda_m)]^2},$$

where $J_0(.)$ and $J_1(.)$ represent the zero-order and first-order Bessel functions respectively and (λ_m ; m = 1, 2….M) are the roots of $J_0(\lambda)$=0.

The FB series coefficients C_m are unique for a given signal, similarly as the Fourier series coefficients are unique for a given signal. However, unlike the sinusoidal basis functions in the Fourier series, the Bessel functions decay over time. This feature of the Bessel functions makes the FB series expansion suitable for non-stationary signals [18, 19, 20].

Thus, the obtained difference time series of EEG signals is subjected to FB series expansion to decompose the signal. In our work, first 500 samples of each difference time series EEG signal were considered and FB coefficients were computed corresponding to these 500 samples.

2.3. Feature Extraction

Feature extraction is the most important step involved in the discrimination process of focal and non-focal bivariate EEG signals. The coefficients obtained from FB decomposition are large in number. Therefore, there is the need to derive efficient features from these coefficients that can be used to accurately distinguish between the two EEG signal classes. In our work, for each signal, we have derived 17 distinguishing features from decomposed series spectrum. These computed features, depending on class being focal or non-focal, uniquely define the characteristics of the spectrum for each class.

These 17 features were selected after extensively experimenting within a large group of theoretical features, as these were able to most distinctly differentiate between the two signal classes.

The coefficients of decomposed bivariate EEG signals were segmented into five smaller parts of equal length. The length of each part was computed experimentally with a length of approximately 60 samples yielding the most efficient features in such a way that first segment consists of (1-60) samples, second segment consists of (61-120) samples, third segment consists of (121-180) samples, fourth segment consists of (121-180) samples, fifth segment consists of (181-240) samples. Features such as first segment peak, second segment peak, third segment peak, fourth segment peak, fifth segment peak and local first peak index, local second peak index, local third peak index, local fourth peak index, local fifth peak index were computed where the first five features represent the coefficient of peak value in the five respective segments and the latter five represent the corresponding local indices of the above peaks.

Further, the whole spectrum is considered for computation of more features such as least squared error, first absolute peak, second absolute peak, third absolute peak and first absolute peak index, second absolute peak index, third absolute peak index where the absolute peaks correspond to first, second and third absolute maximum values of FB coefficients and the absolute peak indices correspond to the indices of these absolute maximums. It is to be noted here that the absolute peak indices are measured from the start of the spectrum in this case. These features are used as input arguments to a binary classifier. Depending on these features, the classifier classifies the signals as focal or non-focal.

2.4 Binary Classification

The techniques used in binary classification are explained in the following subsections.

2.4.1 Support Vector Machine

Support vector machine (SVM) is a binary classifier that works on the principle of supervised learning. For a given dataset with each example marked with the respective class out of the two, an SVM algorithm creates a training model by generating a decision boundary on to a higher dimensional hyper-plane. SVM maps the training examples as points in space, creating a decision boundary such that it has a maximum separation with all the training points, and there is a clear and wide gap between the points corresponding to the two classes.

The basic form of SVM classifier can be expressed as [21, 22]

$$\mathrm{g(x)} = \mathrm{sign}(\sum_{i=1}^{l} \beta_i y_i K(x, x_i) + a)$$

where a, β_i, $K(x, x_i)$ represent the bias, Lagrange multipliers, kernel function respectively.

2.4.2 Least Square Support Vector Machine

The Least square support vector machine (LS-SVM) is a least squares formulation of the SVM classifier. For two-class classification problem in SVM classifier, the discrimination function can be written as follows [23, 24]:

$$u(x) = \mathrm{sign}(\sum_{i=1}^{N} \alpha_i u_i K(x, x_i) + b\ )$$

where u_i, x_i, α_i, b, and $K(x, x_i)$ represent the i^{th} input vector, i^{th} output vector, Lagrange multiplier, bias, and kernel function, respectively.

In this work, we have used the following kernel function together with LS-SVM classifier, which can be expressed as follows:

1. Linear kernel [22]

   $K(x , x_i) = x_i^T x$

2. Polynomial kernel of degree '*d*' [22]

   $K(x , x_i) = (1 + x_i^T x)^d$ where '*d*' is the order of the polynomial.

3. RBF kernel [22]

   $K(x , x_i) = \exp(-||x - x_i||^2/\sigma^2)$ where 'σ' is the RBF kernel parameter.

In our work, we have used SVM classifier and its least squares formulation which is LS-SVM classifier. SVM was implemented with RBF kernel, whereas LS-SVM was implemented with linear, polynomial and RBF kernels.

Chapter 3. Results and Discussion

Decomposition of EEG signals using FB series expansion yields the coefficients whose spectrum can be represented by Figures 3.1 and 3.2 respectively for focal and non-focal EEG signals.

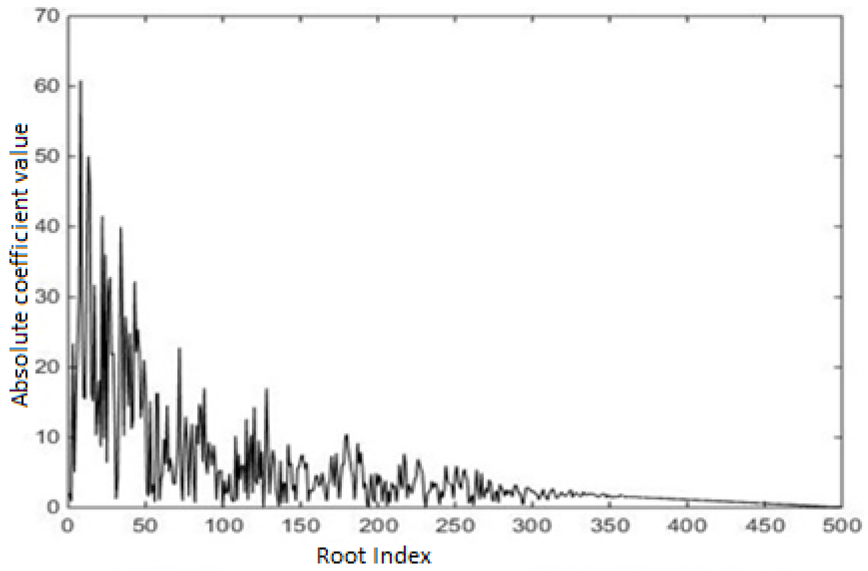

Figure 3.1. Fourier Bessel series spectrum of a difference focal EEG signal

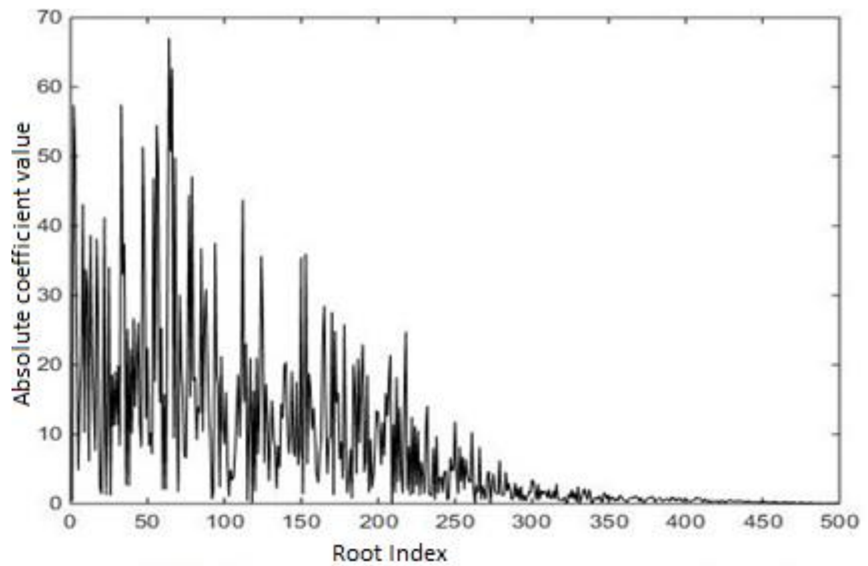

Figure 3.2. Fourier Bessel series spectrum of a difference non-focal EEG signal.

With SVM, RBF kernel was used, whereas with LS-SVM, linear and polynomial kernels were also used along with RBF kernel. Also, training of the classifiers was done using 10-fold cross-validation.

Generally, training of classifiers using the concept of 10-fold cross-validation involves dividing the training dataset into 10 equal parts. Out of the 10 parts, 9 parts are chosen for the training of classifier while the remaining dataset part is chosen for testing on the trained classifier [25]. The classification accuracy is noted.

In our work, this process is continued for nine more iterations where the dataset part for testing is changed, and training dataset is chosen as the rest of the nine parts. The accuracy is noted for each iteration. The final accuracy is the mean of all the accuracies obtained.

Through our methodology, maximum binary classification accuracy was obtained using RBF kernel in LS-SVM, while the least accuracy was obtained using linear kernel. With RBF kernel, most optimum values of σ^2 and γ which yielded the best accuracy were 6.2 and 8 respectively. With polynomial kernel, the best accuracy was obtained for a kernel of degree 3. With linear kernel, the most optimum value of γ was 0.32.

Further, accuracies obtained from classification using SVM and LS-SVM were compared. The maximum classification accuracy obtained using our methodology was 84.25%. Thus, our work and methodology provides a good alternative among other proposed methods for distinguishing focal and non-focal EEG signals. The classification accuracies obtained are summarized in the table below.

Table 3.1. Comparison of classification accuracies for different classifiers with 400 focal and 400 non-focal EEG signals dataset.

Classifier	Accuracy	
SVM	RBF Kernel	76.74%
LS-SVM	RBF Kernel	84.25%
	Polynomial Kernel	78.4%
	Linear Kernel	73.5%

Also, training and testing was done using datasets of size 100, 200, 500 and 800 with equal number of focal and non-focal signals. The results are illustrated in the figure below.

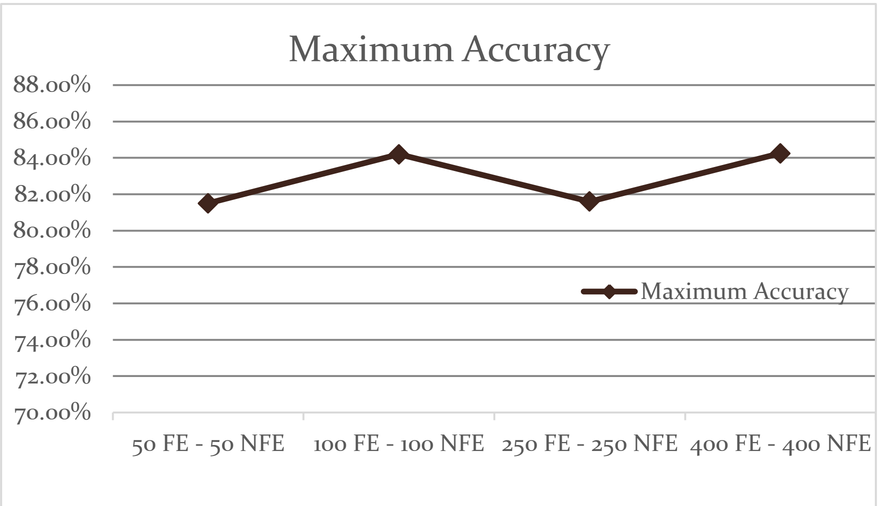

Figure 3.3. Maximum classification accuracies compared for different dataset sizes consisting of focal and non-focal bivariate EEG signals. FE represents focal EEG and NFE represents non-focal EEG signals.

Chapter 4. Conclusion

In our work, we used Fourier-Bessel series expansion as a method of signal decomposition. Features were extracted from the spectrum of these coefficients, which were used in the training and testing of classifiers.

With the two classifiers implemented, results came out better with LS-SVM using RBF kernel. Efficient classification of EEG signals into focal and non-focal category would result in an ease of finding epileptic region of the brain. This would in turn mean, more accurate detection with less time.

With a maximum classification accuracy of 84.25%, our methodology proves to be good alternative among other approaches for efficient detection of focal and non-focal EEG signals.

In the future, we are planning to continue this work, will be introducing new mathematical features with the objective to increase the classification accuracy. Also, a new more efficient classifier will be implemented with these features to further increase the accuracy.

References

[1] Fisher, R.S., Boas, W.V.E., Blume, W., Elger, C., Genton, P., Lee, P. and Engel, J., 2005. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, *46*(4), pp.470-472.

[2] Gloor, P. and Fariello, R.G., 1988. Generalized epilepsy: some of its cellular mechanisms differ from those of focal epilepsy. *Trends in Neurosciences*, *11*(2), pp.63-68.

[3] World Health organization 2017, epilepsy, http://www.who.int/mental_health/neurology/epilepsy/en/index.html (last accessed 6.04.16)

[4] Pati, S. and Alexopoulos, A.V., 2010. Pharmacoresistant epilepsy: From pathogenesis to current and emerging therapies. *Cleve. Clin. J Med.*, *77*(7), pp.457-467.

[5] Acharya, U.R., Sree, S.V., Swapna, G., Martis, R.J. and Suri, J.S., 2013. Automated EEG analysis of epilepsy: A review. *Knowledge-Based Systems*, *45*, pp.147-165.

[6] Bradley, W.G. ed., 2004. *Neurology in clinical practice: principles of diagnosis and management,* Volume 1. Taylor & Francis.

[7] Bhattacharyya, A., Sharma, M., Pachori, R.B., Sircar, P. and Acharya, U.R., 2016. A novel approach for automated detection of focal EEG signals using empirical wavelet transform. *Neural Computing and Applications*, pp.1-11.

[8] Singh, P. and Pachori, R.B., 2017. Classification of focal and nonfocal EEG signals using features derived from Fourier-based rhythms. *Journal of Mechanics in Medicine and Biology*, p.1740002.

[9] Sharma, R., Kumar, M., Pachori, R.B. and Acharya, U.R., 2017. Decision support system for focal EEG signals using tunable-Q wavelet transform. *Journal of Computational Science*, *20*, pp.52-60.

[10] Bhattacharyya, A., Pachori, R.B. and Acharya, U.R., 2017. Tunable-Q wavelet transform based multivariate sub-band fuzzy entropy with application to focal EEG signal analysis. *Entropy*, *19*(3), p.99.

[11] Sharma, M., Dhere, A., Pachori, R.B. and Acharya, U.R., 2017. An automatic detection of focal EEG signals using new class of time–frequency localized orthogonal wavelet filter banks. *Knowledge-Based Systems*, *118*, pp.217-227.

[12] Sharma, R., Pachori, R.B. and Acharya, U.R., 2015. An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures. *Entropy*, *17*(8), pp.5218-5240.

.[13] Das, A.B. and Bhuiyan, M.I.H., 2016. Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain. *Biomedical Signal Processing and Control*, *29*, pp.11-21.

[14] Sharma, R., Pachori, R.B. and Gautam, S., 2014. Empirical mode decomposition based classification of focal and non-focal seizure EEG signals. In *International Conference on Medical Biometrics*, *2014* (pp. 135-140). IEEE

[15] Sharma, R., Pachori, R.B. and Acharya, U.R., 2015. Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals. Entropy, 17(2), pp.669-691.

[16] Andrzejak, R.G., Schindler, K. and Rummel, C., 2012. Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients. *Physical Review E*, *86*(4), p.046206.

[17] Schroeder, J., 1993. Signal processing via Fourier-Bessel series expansion. *Digital Signal Processing*, *3*, pp. 112-124

[18] Pachori, R.B. and Sircar, P., 2006. Speech analysis using Fourier-Bessel expansion and discrete energy separation algorithm. In *Digital Signal Processing Workshop, 12th-Signal Processing Education Workshop*, (pp. 423-428). IEEE.

[19] Pachori, R.B. and Sircar, P., 2007. A new technique to reduce cross terms in the Wigner distribution. *Digital Signal Processing*, 17(2), pp.466-474.

[20] Pachori, R.B. and Sircar, P., 2008. EEG signal analysis using FB expansion and second-order linear TVAR process. *Signal Processing*, *88*(2), pp.415-420.

[21] Cortes, C. and Vapnik, V., 1995. Support-vector networks. *Machine learning*, *20*(3), pp.273-297.

[22] Joshi, V., Pachori, R.B. and Vijesh, A., 2014. Classification of ictal and seizure-free EEG signals using fractional linear prediction. *Biomedical Signal Processing and Control*, *9*, pp.1-5.

[23] Gupta, V., Priya, T., Yadav, A.K., Pachori, R.B. and Acharya, U.R., 2017. Automated detection of focal EEG signals using features extracted from flexible analytic wavelet transform. *Pattern Recognition Letters*, *94,* pp. 180-188.

[24] Suykens, J.A. and Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural Processing Letters*, *9*(3), pp.293-300.

[25] Gupta, V., Bhattacharyya, A. and Pachori, R.B., 2017. Classification of seizure and non-seizure EEG signals based on EMD-TQWT method. In *22nd International Conference on Digital Signal Processing (DSP) 2017,* (pp. 1-5). IEEE.