ISSN 2047-3338

Performance Evaluation of SVM – RBF Kernel for Classifying ECoG Motor Imagery

S. Deepajothi[1] and Dr. S. Selvarajan[2]

***Abstract*– Brain–Computer Interfaces (BCIs) provide a non-muscular channel to communicate with the outside world by means of brain activity. A crucial step for efficient BCI operation is brain signal processing methods. Most BCI systems for humans use scalp recorded electroencephalographic activity, whereas Electrocorticography (ECoG) is a minimally-invasive alternative to electroencephalogram (EEG), providing higher and superior signal characteristics allowing rapid user training and faster communication. Its efficiency is based on brain signal processing methods that classify brain signal patterns in different tasks accurately. Artifacts in raw brain signal make it necessary to pre-process signals for feature extraction. This paper presents a BCI system pre-processing and extracting features from ECoG signals through the use of Symlet Wavelets. Signal classification is done using Support Vector Machine (SVM) with Radial Basis Function (RBF).**

***Index Terms*– Brain–computer interfaces (BCIs), Electrocorticography (ECoG), Symlet Wavelets, Support Vector Machine (SVM) with Radial Basis Function (RBF)**

I. INTRODUCTION

A Brain–Computer Interface (BCI) is a system which allows a person to control special computer applications like a cursor/robotic limb by use of thoughts. The idea was to provide a new communication channel to those paralyzed but was cognitively intact, including those suffering from the locked-in syndrome. BCIs research became active when it grew in the last decade as science improved rapidly [1], [2]. Presently BCI systems for humans use scalp recorded ECG which could have limitations.

Electrocorticography (ECoG) is a minimally-invasive alternative to electroencephalogram (EEG) for BCI systems, which provide superior signal characteristics allowing rapid user training and faster communication. Studies also show that brain regions like the auditory cortex can be trained to control a BCI system using methods similar to those which train the brain's motor regions. This is vital for users with neurological disease, head trauma, or conditions which preclude sensorimotor cortex use for BCI control [3].

In ECoG, neural signals are recorded through disc-like electrodes embedded in a flat strip/grid on the cerebral cortex's [4] subdural surface. ECoG brain signals have advantages when used with BCI systems when compared with EEGs which include increased spatial resolution and signal bandwidth and larger signal amplitude. Hence, independent signals can be differentiated over a range of frequencies - on neighbouring electrodes - using many procedures using both motor and sensory imagery. A few BCI studies with ECoG could provide an optimal balance between signal quality, spatial resolution, temporal resolution, and invasiveness [5], [6]. Leuthardt et al. [7] demonstrated ECoG enables users control a computer cursor in one dimension quickly and accurately.

The control methodology is based on the subject's ability to voluntarily modulate one or more brain rhythms using imagery. Traditionally, motor imagery was used as it was thought to be the most accessible and reliable EEG signal. But, ECoG's advantages enable subjects to learn multiple modalities use - including motor and sensory imagery - to control a BCI application enabling individuals with damaged motor cortex – caused by stroke or other neurological disease - to benefit from BCI systems.

Mental tasks were chosen to activate the brain's various parts enabling easier detection. Sensorimotor activity like body movements/mental imagery including imagining body movement led to oscillatory pattern changes resulting in amplitude suppression known as event related desynchronization. Supervised classification methods recognized such activity patterns to learn data and classes mapping corresponding to mental tasks including left hand movement [8]. From data mining perspective, two reasons made this difficult. First, raw data was noisy and correlated as many electrodes had to be fixed on limited scalp surface with each electrode measuring the activity of thousands of neurons [9]. Selecting an optimal frequency band to extract good features set continued to be an open research problem.

Next, the subject's different degrees of attention and concentration changes affected data quality. Usually, linear classifiers like Fisher's linear discriminant were favoured [8, 10] till recently when many machine learning classifiers were applied including neural networks likes multi-layer perceptrons, probabilistic classifiers, lazy learning classifiers including k-nearest neighbor and state of the art classifiers

[1]Assistant Professor, Department of Computer Science and Engineering, Chettinad College of Engineering and Technology, Karur, Tamilnadu, India

[2]Principal of Muthayammal Technical Campus, Rasipuram, Tamilnadu, India

like support vector machines. But as Lotte et al. (2007) noted, but differing experimental set ups, pre-processing and feature selection reported in various studies ensured that evaluation was tough.

In this study, the Data Set I for the BCI Competition III containing motor imagery in ECoG recordings is used. A BCI system is presented which pre-processes and extracts features from ECoG signals through the use of Symlet Wavelets. Signal classification is done using Support Vector Machine (SVM) with Radial Basis Function (RBF). The performance of the SVM classifier for varying parameters is evaluated.

II. RELATED WORKS

Lotte, et al., [8] reviewed classification algorithms used in EEG BCI systems to identify critical properties. Based on literature, performance was compared, and guidelines provided to choose classification algorithms for specific BCIs. Linear classifiers, neural networks, nonlinear Bayesian classifiers, nearest neighbor classifiers and combinations of classifiers were used in BCI, and they were looked at in detail ultimately proving that while SVM are specifically efficient for synchronous BCI, it performed better due to its regularization property and immunity to the dimensionality curse.

Sherwood et al [11] presented EEG signal classification results for different tasks for support vector machine (SVM) classifiers. EEG being generated from imagined motor, cognitive, and affective tasks. Data had wavelet feature extraction applied on it, performing well even when noise was present, and classifiers were presented with contaminated training data. For six imagined motor tasks and two affective tasks, classifiers achieved more than 80% performance. Cognitive tasks had 70% classification accuracy, the results thereby proving that wavelet features with SVM provided good classification for imagined motor, cognitive and affective tasks.

Lal, et al., [12] proposed feature selection algorithms like Recursive Feature Elimination (RFE) and Zero-Norm Optimization based on SVM training. For feature selection, they provided better solutions compared to the usual filter methods. The algorithms are adapted to select EEG channels. The proposed method revealed that the used channels number could be lowered without increasing motor imagery classification error. Dependent task specific visualization was achieved with the results proving that the suggested method could be used for BCI research where there was no prior knowledge of important channels location.

Barachant et al [13] investigated spatial covariance matrix use as a feature for motor imagery EEG-based classification. A new kernel is derived through the establishment of connection with the Riemannian geometry of symmetric positive definite matrices. Different kernels are tested, combining with SVM on past BCI competition data set demonstrating that this new approach greatly outperformed the state of the art results without spatial filtering. The approach when tested on a BCI competition dataset significantly outperformed conventional CSP method. This kernel can be used in various applications where covariance matrices are main ingredients of feature extraction.

Spuler et al [14] proposed an adaptive SVM-based algorithm and proved through an offline BCI analyzed data recorded with MEG its superiority to other adaptive/non-adaptive classifiers. An online experiment was performed with 8 subjects for verification proving that the proposed adaptive algorithm achieved higher accuracies than non-adaptive baseline algorithms. It was shown that unsupervised online adaption of a SVM classifier in a BCI - for the first time - could significantly increase performance.

Jrad et al [15] focused on sensor weighting to improve classification. An approach An integrating sensor weighting approach in the classification framework was presented. Sensor weights are hyper-parameters to be learned by a SVM. The resulting sensor weighting SVM (sw-SVM) should satisfy a margin criterion, i.e; generalization error. Experimental studies on two data sets - a P300 data set and an Error Related Potential (ErrP) data set – are presented. For the P300 data set (BCI competition III) which has many trails available, the sw-SVM performance is equal with regard to the ensemble SVM strategy which won the competition. For ErrP data set, having limited trials available sw-SVM reveals improved performances when compared to three state-of-the art approaches. Results suggest that sw-SVM can be useful in event-related potentials classification, even with limited training trials.

Ming et al [16] analyzed and classified ERD/ERS response evoked by the left hand, right hand, foot and tongue motor imagery. The signals were spatially filtered by Independent Component Analysis (ICA) before calculating power spectral density (PSD) for related electrodes with SVM being adopted to recognise different imagery patterns based on ERD/ERS feature for signals. The results showed that combining ICA-based signal extraction algorithm and SVM-based classification procedure was effective motor imagery potentials, identification with high and low accuracy rates of 91.4% and 77.6% respectively.

III. MATERIALS AND METHODS

Dataset

Dataset used to evaluate the proposed method is Data Set I for BCI Competition III having motor imagery in ECoG recordings. A subject had to perform imagined movements of left small finger or tongue in a BCI experiment. Electrical brain activity's time series was picked up in these trials. All recordings had a sampling rate of 1000Hz. Recorded potentials were stored as microvolt values after amplification. Each trial included an imagined tongue/imagined finger movement being recorded for 3 seconds. To avoid data reflecting visually evoked potentials, recording intervals started 0.5 seconds after the end of the visual round. Brain activity for 278 trials was considered training data and similar activity for100 trials was considered test data [17].

Symlet Wavelets

EEG signals are typical non-stationary, random signals, rich in weak signal frequency components. Wavelet transforms analyses transient and time-varying characteristics of non-stationary signals like EEG and ECoG. Wavelet is a waveform bound in both frequency and time. Fourier analysis includes breaking a signal into various frequency sine waves. Similarly, wavelet analysis breaks up a signal into the mother wavelet's shifted and scaled versions. A continuous wavelet transform (CWT) is the sum over all time of a signal multiplied by scaled and shifted versions of wavelet function ψ.

Mathematically the continuous wavelet is defined by:

CWT results are many wavelet coefficients C, a function of scale and position. Multiplying each coefficient by scaled and shifted wavelets yield the original signal's constituent wavelets. Symlets are nearly symmetrical wavelets proposed by Daubechies as modifications to the db family [18] and

$$C(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(scale, position, t)dt$$

properties of both wavelet families are similar. Symlets is compactly supported wavelets which for a given support width have least asymmetry and the highest number of vanishing moments.

Support Vector Machine (SVM)

Given a set of features that can be represented in space, SVM maps features non-linearly into n dimensional feature space when provided with features set that can be represented in space. When a kernel is introduced with high computation the algorithm uses inputs as scalar products with classification being solved by translating the issue into a convex quadratic optimization problem with a clear solution being obtained by convexity [19]. In SVM, an attribute is a predictor variable and a feature a transformed attribute. A set of features describing an example is a vector. Features define the hyperplane. SVM aims to locate an optimal hyperplane separating vector clusters with a class of attributes on one side of the plane with the on the other side. The margin is the distance between hyperplane and support vectors. SVM analysis orients the margin that space between it and support vectors is maximized. Figure 1 shows a simplified SVM process overview.

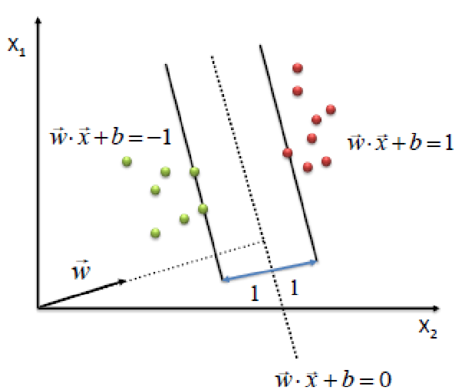

Fig. 1: Support vector machine

Given a training set of $(x_i, y_i), i1, 2, \ldots. l$ where $x_i \in R^n$ and $y \in \{1, -1\}^l$, SVM has to solve the optimization problem [15] of: $\min_{w,b,\xi} \frac{1}{2} w^T w + C\sum_{i=1}^{l} \xi_i$

Subject to $y_i\left(w^T\phi(x_i)+b\right) \geq 1-\xi_i$ and $\xi_i \geq 0$.

The function ϕ maps the vectors x_i in higher dimensional space. C>0 is penalty parameter of the error term.

This optimization model is solved through the use of the Lagrangian method, equal to the method for solving optimization problems in a separable case. One maximizes the dual variables Lagrangian:

$$\underset{\alpha}{Max} \;\; L_D(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{m} \alpha_i\alpha_j y_i y_j \langle x_i . x_j \rangle$$

subject to: $0 \leq \alpha_i \leq C \;\; i = 1,..,m$ and $\sum_{i=1}^{m} \alpha_i y_i = 0$

To find the optimal hyperplane, a dual LagrangianLD(α) should be maximized as regards non-negative α_i under the constrains $\sum_{i=1}^{m} \alpha_i y_i = 0$ and $0 \leq \alpha_i \leq C$. The penalty parameter C, now the upper bound on α_i, is user determined.

A kernel function is defined as $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$. The Radial Basis function is given as follows:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \;, \gamma > 0$$

A proper parameter setting improves SVM classification accuracy. There are two parameters to be determined in the SVM model with the RBF kernel: C and gamma (c). Instinctively the gamma parameter defines the distance a single training example can reach, with low values meaning 'far' and high values meaning 'close'. The C parameter trades off training examples misclassification against decision surface simplicity. A low C ensures a smooth decision surface while a high C attempts to classify training examples correctly. Experiments are undertaken to evaluate SVM performance through variations of the Gamma and C parameters.

IV. RESULTS AND DISCUSSION

Experiments were conducted using 278 numbers of instances from the dataset. Features were extracted using Symlet wavelet. 193 attributes were used to classify the instances. All the experiments were conducted for 10-fold cross validation. Two sets of experiments were conducted: In set 1, the Gamma value is maintained constant at 0.125 and the C value is varied (0.125, 0.5, 0.75 and 1) and in set II, the C value is maintained at a constant value of 0.125 and Gamma value is varied (0.25, 0.5, 0.75 and 1). The classification accuracy and the root mean square error

(RMSE) achieved is tabulated in Table 1. Figure 2 shows the classification accuracy and Figure 3 show the RMSE.

Table 1: Classification Accuracy and RMSE for varying Gamma and C value

SVM RBF Parameters	Classification Accuracy %	RMSE
Gamma, C		
0.125, 0.125	77.3381	0.476
0.125, 0.5	77.3381	0.476
0.125, 0.75	77.3381	0.476
0.125, 1	77.3381	0.476
0.25, 0.125	71.9424	0.5297
0.5, 0.125	68.3453	0.5626
0.75, 0.125	63.3094	0.6057
1, 0.125	62.9496	0.6087

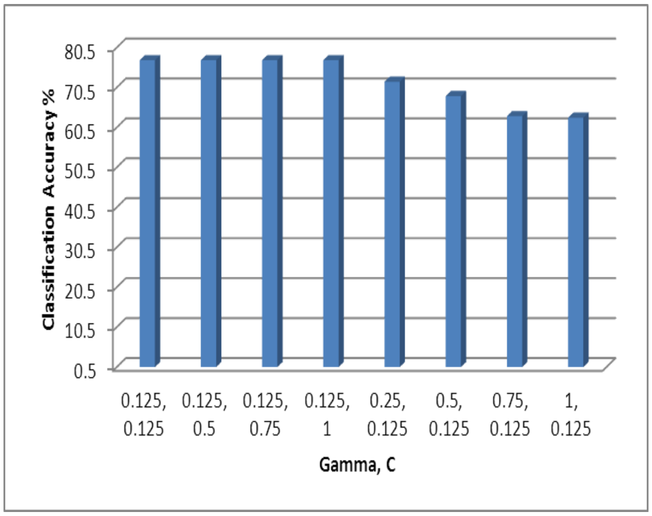

Fig. 2: Classification Accuracy

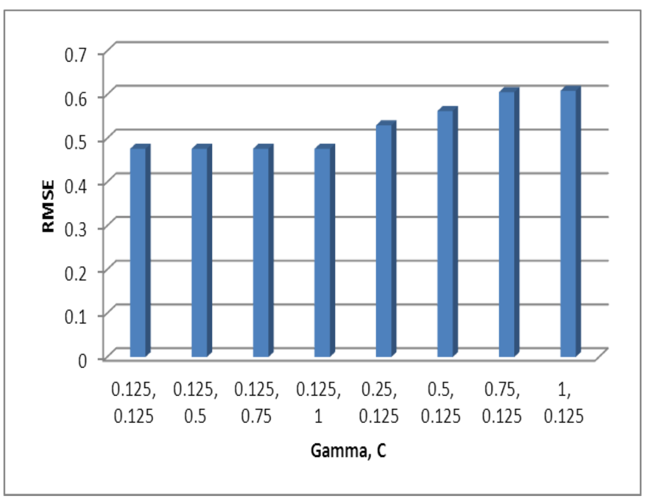

Fig. 3: Root Mean Square Error

It is observed from the Table and Figures that the varying of the parameter C has no effect on the classification accuracy or the RMSE. Also, higher value of Gamma leads to inefficient performance of the SVM. The best classification accuracy was obtained for Gamma value of 0.125. Table 2 tabulates the precision and recall achieved.

Table 2: Precision and Recall

SVM RBF Parameters	Precision	Recall
Gamma, C		
0.125, 0.125	0.773	0.773
0.125, 0.5	0.773	0.773
0.125, 0.75	0.773	0.773
0.125, 1	0.773	0.773
0.25, 0.125	0.719	0.719
0.5, 0.125	0.683	0.683
0.75, 0.125	0.633	0.631
1, 0.125	0.629	0.626

Similar to the classification accuracy, precision and recall are high when Gamma value is 0.125. Further investigations are required to improve the classification of the ECoG signals.

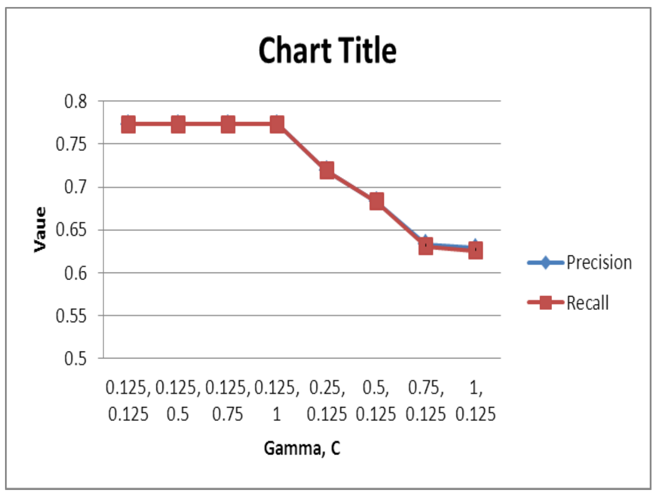

Fig. 4: Precision and Recall

V. CONCLUSION

In this study, a BCI system is presented which pre-processes and extracts features from ECoG signals through the use of Symlet Wavelets. Signal classification is done using Support Vector Machine (SVM) with Radial Basis Function (RBF). Data Set I for the BCI Competition III containing motor imagery in ECoG recordings is used for evaluating the system. The performance of the SVM classifier for varying parameters is evaluated. Experiments were carried out through tenfold cross validation and the accuracy achieved is comparable with results from other research in the literature but further work is required to improve classification accuracy.

REFERENCES

[1] Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain–computer interfaces for communication and control. Clin Neurophysiol 2002;113:767–91.

[2] Dornhege, G., Millán, J. D. R., Hinterberger, T., McFarland, D., & Müller, K. R. (2007). Toward brain-computer interfacing (Vol. 74). Cambridge, MA: MIT press.

[3] Wilson, J. A., Felton, E. A., Garell, P. C., Schalk, G., & Williams, J. C. (2006). ECoG factors underlying multimodal control of a brain-computer interface. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 14(2), 246-250.

[4] Felton, E. A., Wilson, J. A., Williams, J. C., & Garell, P. C. (2007). Electrocorticographically controlled brain-computer interfaces using motor and sensory imagery in patients with temporary subdural electrode implants. Journal of neurosurgery, 106(3), 495-500.

[5] Felton EA, Wilson JA, Radwin RG, Williams JC, Garell PC: Electrocorticogram-controlled brain-computer interfaces in patients with temporary subdural electrode implants. Neurosurgery 57: 425, 2005

[6] Liu, C., Zhao, H. B., Li, C. S., & Wang, H. (2010, October). Classification of ECoG motor imagery tasks based on CSP and SVM. In Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on (Vol. 2, pp. 804-807). IEEE.

[7] Leuthardt EC, Schalk G, Wolpaw JR, Ojemann JG, Moran DW: A brain-computer interface using electrocorticographic signals in humans. J Neural Eng 1:63–71, 2004

[8] Lotte F, Congedo M, Lecuyer A, Lamarche F and Arnaldi B (2007) “A review of classification algorithms for EEG-based brain–computer interfaces” J. Neural Eng. PP: 1-13.

[9] Thulasidas, M., Guan, C., & Wu, J. (2006). Robust classification of EEG signal for brain-computer interface. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 14(1), 24-29.

[10] Müller, K. R., Krauledat, M., Dornhege, G., Curio, G., & Blankertz, B. (2004). Machine learning techniques for brain-computer interfaces.

[11] Sherwood, J., and Derakhshani, R. 2009. “On classifiability of wavelet features for EEG based brain-computer interfaces.” International Joint Conference on Neural Networks, Atlanta, GA, June 2009, PP: 2895-2902.

[12] Lal TN, Schroder M, Hinterberger T, Weston J, Bogdan M, Birbaumer N, Scholkopf B. Support Vector Channel Selection in BCI. IEEE Trans Biomed Eng. 2004;51(6):1003–1010.

[13] Barachant, A., Bonnet, S., Congedo, M., & Jutten, C. (2012, April). BCI Signal Classification using a Riemannian-based kernel. In Proceeding of the 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (pp. 97-102).

[14] Spüler, M., Rosenstiel, W., & Bogdan, M. (2012). Adaptive SVM-based classification increases performance of a MEG-based Brain-Computer Interface (BCI). Artificial Neural Networks and Machine Learning–ICANN 2012, 669-676.

[15] Jrad, N., Congedo, M., Phlypo, R., Rousseau, S., Flamary, R., Yger, F., & Rakotomamonjy, A. (2011). sw-SVM: sensor weighting support vector machines for EEG-based brain–computer interfaces. Journal of neural engineering, 8(5), 056004.

[16] Ming, D., Sun, C., Cheng, L., Bai, Y., Liu, X., An, X., ... & Luk, K. D. K. (2010, September). ICA-SVM combination algorithm for identification of motor imagery potentials. In Computational Intelligence for Measurement Systems and Applications (CIMSA), 2010 IEEE International Conference on (pp. 92-96). IEEE.

[17] Thomas Lal, Thilo Hinterberger, Guido Widman, Michael Schröder, Jeremy Hill, Wolfgang Rosenstiel, Christian Elger, Bernhard Schölkopf, Niels Birbaumer. Methods Towards Invasive Human Brain Computer Interfaces. Advances in Neural Information Processing Systems

[18] Vinay U. Kale & Nikkoo N.Khalsa "Performance Evaluation of Various Wavelets for Image Compression of Natural and Artificial Images" International Journal of Computer Science & Communication,Vol. 1, No. 1, January-June 2010, pp. 179-184

[19] B. E. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. In Proceedings of The Fifth Annual Workshop on Computational Learning Theory, pages 144-152. ACM Press, 1992.