Proceedings of the 2011 IEEE/ICME
International Conference on Complex Medical Engineering
May 22 - 25, Harbin, China

Developing a Logistic Regression Model with Cross-Correlation for Motor Imagery Signal Recognition

Siuly and Yan Li
Department of Mathematics and Computing, University of Southern Queensland Toowoomba, QLD 4350, Australia
{siuly, liyan}@usq.edu.au

Jinglong Wu and Jingjing Yang
Biomedical Engineering Laboratory Graduate School of Natural Science and Technology Okayama University, Japan
wu@mech.okayama-u.ac.jp

***Abstract---Classification* of motor imagery (MI)-based electroencephalogram (EEG) signals is a key issue for the development of brain-computer interface (BCI) systems. The objective of this study is to develop an algorithm that can distinguish two categories of MI EEG signals. In this paper, we propose a new classification algorithm for two-class MI signals recognition in BCIs. The proposed scheme develops a novel cross-correlation-based feature extractor, which is aided with a logistic regression model. The present method is tested on dataset IVa of BCI Competition III, which contain two-class MI data for five subjects. The performance is objectively computed using a *k*-fold cross validation (*k=10*) method on the testing set for each subject. The results of this study are compared with the recently reported eight methods in the literature. The results demonstrate that our proposed method outperforms the eight methods in terms of the average classification accuracy.**

Index Terms---Electroencephalogram (EEG), Brain-computer interface (BCI), Motor imagery (MI), Cross-correlation technique and Logistic regression model.

I. INTRODUCTION

Over the last decades, brain-computer interface (BCI) has emerged as a new frontier in assistive technology since it provides an alternative communication channel between a user's brain and the outside world [1]. A BCI system allows individuals with motor disabilities to control objects in their environments (such as a light switch in their room or television, wheelchair, neural prosthesis and computer) using their brain signals only [1, 2]. This could be accomplished by measuring specific features of the user's brain activity that relates to his/her intention to perform the control. The ultimate object of a BCI is to provide humans an alternative communication channel allowing direct transmission of messages from the brain by analysing the brain's mental activities [3, 4].

In recent years, increasing attention has been devoted to the analysis of EEG signals in motor imagery (MI) problems related to BCI applications as MI-based BCI provides a promising communication channel for the patients who suffer from motor disabilities [5]. Motor imagery (MI) may be seen as the mental rehearsal of a motor act, such as movements of hands, foots, fingers or tongue without any overt motor activity [6,7]. As a BCI system works through the EEG brain signals, a big challenge, therefore, is for the BCI systems to correctly and efficiently identify different EEG signals of different MI tasks using appropriate classification algorithms. In order to control a BCI system, the user must produce different brain activity patterns that will be identified by the system and translated into commands [8]. In most existing BCIs, this identification relies on a classification algorithm i.e. an algorithm that aims at automatically estimating the class of data as represented by a feature vector [9].

Due to the rapidly growing interest in the MI-based BCIs, a number of methods are reported by different researchers for the MI EEG signal classification. Most recently reported eight methods are discussed here and those methods were implemented on dataset IVa of BCI Competition III.

Yong et al. [10] reported a sparse spatial filter optimization for EEG channel reduction in BCI where the spatial filter was used to project the signals and the variance of the projected signals was the only feature used in the linear discriminant analysis (LDA) as the input for the classification. They achieved a classification accuracy of 57.5% for subject '**aa**'**,** 86.9% for subject '**al**', 54.4% for subject '**av**', 84.4% for subject '**aw**' and 84.3% for subject '**ay**', and the average accuracy was 73.5%. But the major limitation of that study was that they manually selected their regularization parameter.

Lu et al. [11] introduced a regularized common spatial patterns (R-CSP) algorithm by incorporating the principle of generating learning for EEG signal classification. The reported classification accuracy rates were 69.6%, 83.9%, 64.3%, 70.5%, and 82.5% for subjects '**aa**', '**al**', '**av**', '**aw**' and '**ay**', respectively and the average accuracy rate was 74.2% for all subjects. It was reported that the algorithm was effective in a small sample setting.

Lotte et al. [12] proposed four methods representing a family of a theoretical framework based on regularized common spatial patterns (RCSP). Their proposed methods are regularized CSP with selected subjects (SSRCSP) [12], CSP with Tikhonov regularization (TRCSP) [12], CSP with weighted Tikhonov regularization (WTRCSP) [12] and spatially regularization (SRCSP) [12]. The average classification success rate reached at 73.56% for the SSRCSP, 77.98% for the TRCSP, 75.93% for the WTRCSP and 78.63% for the SRCSP. The experimental results showed that the TRCSP and WTRCSP algorithms were better than other two algorithms. Their algorithms were based on common spatial patterns (CSP) method, although the CSP is a popular method in BCI applications, but it is very sensitive to noise, and often over-fit with small training sets.

Lu et al. [13] introduced a regularization and aggregation technique with CSP for EEG signal classification in a small sample setting (SSS). The obtained classification accuracy rates were 76.8%, 98.2%, 74.5%, 92.9% and 77.0% for subjects '**aa**', '**al**', '**av**', '**aw**' and '**ay**', respectively, for experiment III. The overall accuracy performance was 83.9%. The main drawback is that this method is only applicable for a small sample setting.

Most recently, Siuly et al. [14] reported a clustering technique-based least square support vector machine algorithm (LS-SVM) for EEG signal classification. It achieved the classification accuracy of 92.63% for subject '**aa**', 84.99% for subject '**al**', 90.77% for subject '**av**', 86.50% for subject '**aw**' and 86.73% for subject '**ay**'. The average accuracy performance was 88.32%. It is known that the parameters of the LS-SVM method can significantly affect the classification performance but the study in [14] did not select the parameters optimally through any technique. They manually selected the parameters for the LS-SVM method.

Most of the reported methods are limited in their success and effective only in a small sample setting. In most of the cases, the methods did not select their parameters using a suitable technique while the parameters significantly affect the classification performance. To overcome these problems, this paper proposes a new approach which can discriminate two-class MI tasks for the development of BCI systems. In the proposed algorithm, a cross-correlation technique is developed for feature extraction and a logistic regression is applied to classify the obtained features. To the best of our knowledge, the cross-correlation technique and the logistic regression model have not been used together before for the MI task recognition in BCI applications.

In this study, the experimental evaluation is performed on a benchmark dataset IVa of BCI Completion III. The dataset of the MI EEG signals is firstly processed by using the cross-correlation technique and then six statistical features (discussed in Section II (A)) are computed from a resultant data to create the feature vectors. Such vectors serve as the inputs to a logistic regression model that provides the final decisions. In order to evaluate the performance of the proposed approach, specificity, sensitivity and classification accuracy are calculated through a *k*-fold-cross-validation method (*k*=10). Then the experimental results of the proposed method are compared with recently reported eight methods. The research outcomes demonstrate that the present method is superior to the eight other methods in terms of the average classification accuracy.

The remainder of this paper is organized as follows: Section II presents the description of the proposed methodology. Section III provides the experimental set up. The experimental dataset, implementation procedure and results are discussed in this section. This section also presents a comparative study of the proposed method to the recent reported eight methods. Finally Section IV draws the conclusions of the paper and gives an idea about our future work.

II. PRPOSED METHOD

This paper develops a powerful method for the MI EEG signal classification in the development of BCI systems, as illustrated in Fig.1. This method provides an important framework to classify two-class MI based-EEG signals from the recorded EEG signals. As shown in Fig. 1, the building components of the proposed system have the following tasks: the motor imagery of a subject produces brain signals (e.g. EEG signals) from the brain. The 'EEG feature extraction by cross-correlation' block transforms the resultants signals into feature values that carry the characteristics of the original patterns.

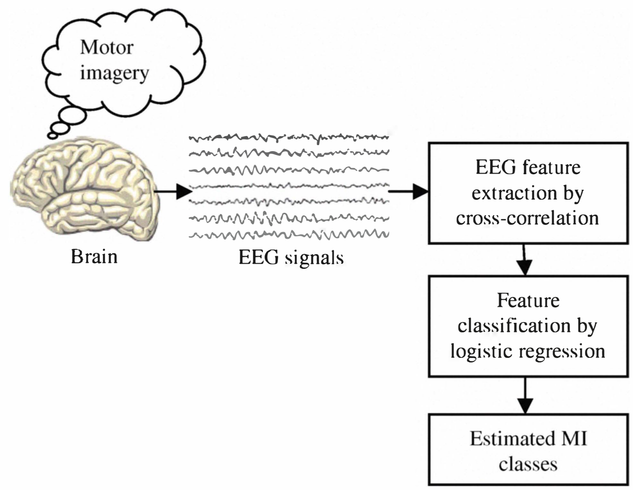

Fig. 1 Schematic diagram of the proposed method for the MI EEG signal classification in BCIs.

The 'feature classification by logistic regression' block performs the classification task using the features obtained from the cross-correlation as the inputs to a logistic regression classifier. The classification results are obtained in the 'estimated MI classes' block. The feature extraction and the classification procedures of the proposed approach are discussed in the following two consecutive sections.

A. Feature Extraction by Cross-Correlation Technique

This study employs a cross-correlation technique for feature extraction from the original MI EEG data. This technique measures the degree of similarity of two signals as a function of a time-lag applied to one of them [15, 16]. The similarity of two signals is numerically evaluated by summing the products of identical time samples of each signal, as shown in (1). The cross-correlation of two finite signals, $x[i]$ and $y[i]$, each of n (n>1) sample points, is defined as [17]

$$R_{xy}[m] = \sum_{i=0}^{n-|m|-1} x[i]y[i-m];\ m = -(n-1)..0,1,2,..(n-1) \quad (1)$$

Here, m represents time-shift parameters known as lag and $R_{xy}[m]$ is a cross-correlated sequence. As each of the signals,

$x[i]$ and $y[i]$, consists of n finite number of samples, the resultant cross-correlation sequence has ($2n$-1) samples. If $x[i]$ and $y[i]$ are not the same length, for example, $x[i]$ and $y[i]$ have n and m samples, respectively, and if $n>m$, the resultant cross-correlation sequence has ($2n$-1) samples. The shorter vector, say $y[i]$, is zero-padded to the length of the longer vector, here $x[i]$. In this study, the features extraction through the cross-correlation technique follows the following steps:

Step 1: a reference signal is selected from any of the two classes MI signals. A signal of any class can be considered as a reference signal as there are no specific requirements for selecting a reference signal in the cross-correlation analysis. In (1), $x[i]$ is considered as the reference signal and $y[i]$ is regarded as any other non-reference signal in one subject of the two-class MI EEG data.

Step 2: The cross-correlation sequence, denoted by '$R_{xy}[m]$', is calculated using a reference signal and any other non-reference signal using the cross-correlation technique. In this study, (1) is used to compute a cross-correlation sequence. The graphical presentation of a cross-correlation sequence is called a cross-correlogram. The reference signal of a class is cross-correlated with the data of the remaining signals of this class and the data of all signals of another class.

Step 3: To reduce the dimensions of each cross-correlation sequence, this study considers six statistical features such as *mean, median, mode, standard deviation, maximum* and *minimum* values, which are regarded as a condensed representation, ideally containing all important information of the original patterns [18]. These features, which are used as the inputs (independent variables) to the logistic regression model for the classification, are calculated from each cross-correlation sequence or cross-correlogram.

B. Classification by Logistic Regression Model

In this paper, a logistic regression model is used as a classifier to classify the two categories MI EEG features obtained by the cross-correlation technique. The logistic regression model is developed to predict the probability of the two-class MI tasks considered as a dependent variable using six independent variables.

Suppose we have a data set, where $x_1, x_2, \ldots\ldots, x_n$ are a vector of input features and y is its class label either 0 or 1. Here $x_1, x_2, \ldots\ldots, x_n$ are treated as independent/predictor variables and y is a dependent variable. Under the logistic regression framework, the probability of the dependent variable y, when y belongs to class 1, is defined as [19, 20]

$$P(y=1|\, x_1, x_2, \ldots\ldots, x_n) = \pi = \frac{e^{\beta_0 + \sum_{i=1}^{n} \beta_i x_i}}{1 + e^{\beta_0 + \sum_{i=1}^{n} \beta_i x_i}} \qquad (2)$$

Here, π is a conditional probability of the form $P(y=1|x_1, x_2, \ldots\ldots, x_n)$. On the other hand, the probability of y, when y belongs to class 0, can be calculated as $1-\pi = 1-P(y=1|x_1, x_2, \ldots\ldots, x_n) = P(y=0|\, x_1, x_2, \ldots\ldots, x_n)$. In (2), β_0 is an intercept and $\beta_1, \beta_2, \ldots, \beta_n$ are the regression coefficient related to the independent variables $x_1, x_2, \ldots. x_n$. These parameters are estimated by maximum likelihood estimation (MLE) technique. Note that in logistic regression, a linear relationship between the independent and dependent variables is not assumed in general, nor does it require normally distributed independent variables [21].

In this research, we consider the MI as a dependent variable, termed y and the six statistical features are treated as six independent variables (n=6). The six independent variables used in (2) are x_1=*mean* values, x_2=*maximum* values, x_3=*minimum* values, x_4=*standard deviation* values, x_5=*median* values and x_6=*mode* values. It is known that the dependent variable y, has two values, 0 or 1, in the logistic regression. For dataset IVa in brain-computer interface (BCI) Competition III [22], the RH MI class is treated as 0 and the RF MI class as 1.

III. EXPERIMENTAL SET UP

A. Data Description

A publically available benchmark dataset IVa of BCI Competition III is used to evaluate the efficacy of the proposed approach in this paper. This dataset contains EEG records, which were collected during MI tasks.

Dataset IVa [22, 23] in BCI Competition III was recorded from five healthy subjects (labelled '**aa**', '**al**', '**av**', '**aw**', '**ay**') and each of the five subjects performed two MI tasks denoted as two classes: right hand (denoted by 'RH') and right foot (denoted by 'RF').The subjects sat in comfortable chairs with their arms resting on armrests. This data set contains MI EEG data from the four initial sessions without feedback. The EEG signals were recorded from 118 electrodes according to the international 10/20-system. There were 280 trials for each subject, namely 140 trials for each task per subject. During each trial, the subject was required to perform either of the two (RH and RF) MI tasks for 3.5 seconds. This study uses the down-sampled data at 100 Hz where the original sampling rate is 1000 Hz.

B. Implementation of the Proposed Method

This section discusses how the proposed method is implemented on dataset IVa of BCI Competition III. Table I presents the information of the original data for dataset IVa of BCI Competition III.

TABLE I
THE INFORMATION OF ORIGINAL DATA FOR DATASET IVa OF BCI III

subject	Dataset IVa, BCI Competition III		
	Sizes of data (RH and RF)	Among 280 trials	
		Number of training trials (labelled)	Number of testing trials (unlabelled)
aa	298458×118	168	112
al	283574×118	224	56
av	283042×118	84	196
aw	282838×118	56	224
ay	283562×118	28	252

It is seen from Table I that every sample of the training trials contains class labels but the testing trials do not have class labels to the samples. In this paper, we used the training trials in our experiments as the proposed algorithm requires a class label at each data point.

As the MI EEG signals are naturally highly subject-specific depending on physical and mental tasks, each subject is considered separately for an experiment in this study. Fig. 2 shows the typical EEG signals of the RH and RF MI classes of subjects **'aa'**, **'al'**, **'av'**, **'aw'** and **'ay'**. As mentioned before, there are 118 signals in each of the two classes of a subject. In the case of each subject, the signal of Fp1 channel of the RH MI class is considered as a reference signal.

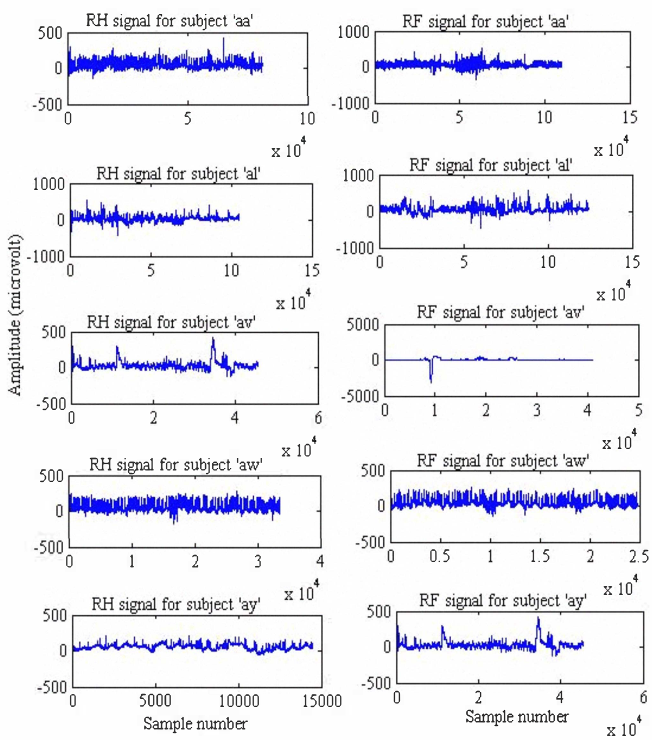

Fig. 2 Typical EEG signals of the RH and RF MI classes from each subject.

For example, in subject **'aa'**, the signal of Fp1 channel of the RH MI class is a reference signal and this reference signal is cross-correlated with the data from the remaining 117 signals of the RH MI class. In the RF MI class of subject **'aa'**, this reference signal is also cross-correlated with the data of all 118 signals of this class. Thus, for subject **'aa'**, a total of 117 cross-correlation sequences/cross-correlograms are obtained from the RH MI class and 118 from the RF MI class. The same process is followed for subjects **'al'**, **'av'**, **'aw'** and **'ay'**, in this study.

Fig. 3 depicts typical results of cross-correlation sequences called cross-correlograms for the RH and the RF MI data of subjects **'aa'**, **'al'**, **'av'**, **'aw'** and **'ay'**, respectively. The cross-correlation sequences (R_{xy}) are obtained using (1) for each lag (m). From each cross-correlation sequence/cross-correlogram, a set of six statistical features, such as *mean, median, mode, standard deviation, maximum* and *minimum* are calculated. Thus, in the case of each subject, we obtain 117 feature vectors of 6 dimensions for the RH class and 118 feature vectors of the same dimensions for the RF class. Finally we obtain a total of 235 feature vectors of 6 dimensions from the two-class MI EEG signals of a subject.

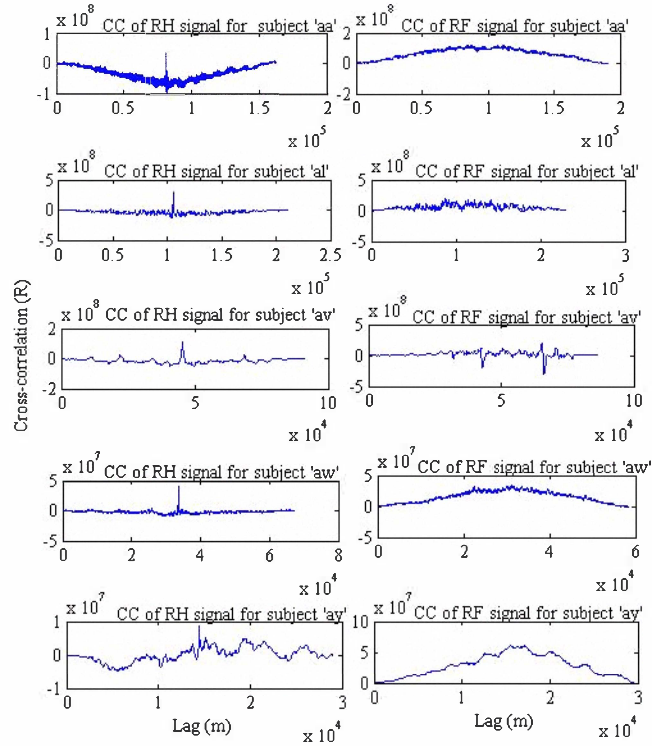

Fig. 3 Typical cross-correlograms (denoted by CC) of the RH and RF MI classes for each subject.

This study employs a 10-fold cross-validation procedure [14] for the performance evaluation of the proposed method. According to the 10-fold cross-validation process, the feature vector set of each subject is randomly divided into 10 subsets with approximately equal sizes and the procedure is repeated 10 times (the folds). Each time, one subset is used as a testing set and the remaining nine subsets are used as a training set. The obtained classification accuracy of each of 10 times on the testing set is averaged, called 10-fold cross-validation accuracy. In addition, specificity and sensitivity of each of the 10 folds are calculated for the testing set and then the 10 values of the specificity and 10 values of the sensitivity are averaged, separately. In each of the 10-folds, specificity, sensitivity and accuracy are calculated using the following formulas [24, 25, 26]:

- Specificity: number of correct classified RH MI segments/number of total RH MI segments
- Sensitivity: number of correct classified RF MI segments/number of total RF MI segments
- Accuracy: number of correct classified segments/number of total segments

It is observed from the 10-fold cross-validation that in each trial, the data are divided into two groups, the training and the testing sets. In this paper, the training sets are applied to train the classifier and the testing vectors are used to verify the effectiveness of the classifier.

C. *Results with Discussion*

The logistic regression model in (2) utilizes the training and testing sets as the inputs separately for estimating the probability of the MI tasks (dependent variable y). All experimental results for dataset IVa are presented based on the testing set in this paper. From the results of the logistic regression model, it can be reviewed on how many values of each class are predicted correctly and how many are predicted incorrectly for the dependent variable y.

Fig. 4 provides a detailed information about the correct classification accuracy of the proposed method for each of the 10 folds of each subject. Fig. 4 is plotted to display the individual correct classification accuracy rate against each of the 10-folds by the 10-fold cross-validation method for subjects **'aa'**, **'al'**, **'av'**, **'aw'** and **'ay'**, respectively.

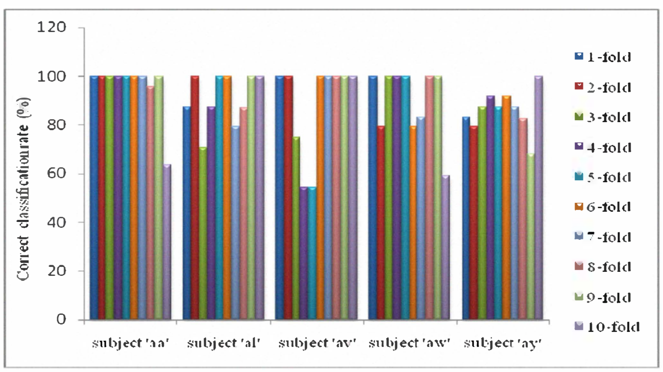

Fig. 4 Correct classification accuracy rate of each of the 10 folds in each subject.

From the figure, it is observed that, in most of the cases, the proposed approach yields a high classification performance for each of the 10-folds of each subject, which is close to 100%.

Table II presents the classification results of the logistic regression classifier for five subjects of dataset IVa. In Table II, the results of each subject are reported in terms of mean ± standard deviation of specificity, sensitivity and accuracy over a 10-fold cross-validation procedure on the testing set. The proposed method achieves the specificity (RH MI success rates) of 94.54%, 88.17%, 88.34%, 87.95%, 85.37% for subjects **'aa'**, **'al'**, **'av'**, **'aw'** and **'ay'**, respectively; and the sensitivity values (RF MI success rates) are 97.27% for subject **'aa'**, 94.17% for subject **'al'**, 88.34% for subject **'av'**, 92.19% for subject **'aw'** and 86.44% for subject **'ay'** using the 10-fold cross-validation method. The table shows that the proposed logistic regression with the cross-correlation features produces an accuracy of 95.93% for subject **'aa'**, 91.20% for subject **'al'**, 88.34% for subject **'av'**, 90.08% for subject **'aw'** and 85.92% for subject **'ay'.** Table II also reports that the standard deviations of each subject for the proposed approach are low, which indicates the reliability of the proposed method.

TABLE II
CLASSIFICATION RESULTS BY THE 10-FOLD CROSS-VALIDATION FOR THE DATASET IVa

subject	Mean ± standard deviation (%)		
	Specificity	Sensitivity	Accuracy
aa	94.54±14.36	97.27±8.63	95.93±11.44
al	88.17±14.35	94.17±8.83	91.20±10.50
av	88.34±19.73	88.34±19.73	88.34±19.61
aw	87.95±16.64	92.19±12.21	90.08±14.27
ay	85.37±9.65	86.44±8.46	85.92±8.49
Average	88.87±3.39	91.68±4.37	90.29±3.73

As seen from the table the proposed technique provides an average specificity of 88.87%, average sensitivity of 91.68% and average classification accuracy of 90.29%. The average standard deviation for specificity, sensitivity and accuracy are not high, which are 3.39, 4.37 and 3.73, respectively, that reflects the consistency of the present algorithm.

C. *Comparisons with Recent Reported Eight Methods*

In order to examine the efficiency of the proposed algorithm, this section provides the comparisons of our approach with eight other recently reported techniques. Those eight algorithms are already discussed in Section I. Table III reports the comparison results of the classification accuracy rates for the proposed method and the eight algorithms for dataset IVa. This table shows the classification performance for the five subjects as well as the overall mean accuracy values.

From Table III, it is noted that the proposed cross-correlation-based logistic regression algorithm provides the highest classification rate of 95.93% for subject **'aa'** among the eight other algorithms.

TABLE III
PERFORMANCE COMPARISONS OF THE PROPOSED METHOD WITH EIGHT OTHER METHODS IN THE LITERATURE

Method	Classification accuracy rate (%)					
	aa	**al**	**av**	**aw**	**ay**	Avg[i]
Prop. Method[*]	95.93	91.20	88.34	90.08	85.92	90.29
CT-LS-SVM[a] [14]	92.63	84.99	90.77	86.50	86.73	88.32
R-CSP-A[b] [13]	76.80	98.2	74.5	92.90	77.0	83.90
SSRCSP[c] [12]	70.54	96.43	53.57	71.88	75.39	73.56
TRCSP[d] [12]	71.43	96.43	63.27	71.88	86.9	77.98
WTRCSP[e] [12]	69.64	98.21	54.59	71.88	85.32	75.93
SRCSP[f] [12]	72.32	96.43	60.2	77.68	86.51	78.63
R-CSP-GL[g] [11]	69.6	83.9	64.3	70.5	82.5	74.20
SSFO[h] [10]	57.5	86.9	54.4	84.4	84.3	73.50

[*]=Proposed method
[a]=Clustering technique-based least square support vector machine;
[b]=Regulized common spatial pattern with aggregation;
[c]=CSP with selected subject;
[d] =CSP with Tikhonov regularization;
[e]=CSP with weighted Tikhonov regularization;
[f] =Spatially regularized CSP;
[g]= Regulized common spatial pattern with generic learning;
[h] =Sparse spatial filter optimization;
[i]=Average;

With the proposed method, we obtain the accuracy rates of 91.20% for subject **'al'**, 88.34% for subject **'av'**, 90.08 for subject '**aw**' and 85.92% for subject '**ay**'. It is observed from the table that the classification accuracy rates of the proposed method of subjects '**al**', '**av**', '**aw**' and '**ay**' are a bit less than the WTRCSP [12], CT-LS-SVM [14], R-CSP-A [13] and TRCSP [12], respectively.

Further looking at the performance comparison in Table III, it is noted that our proposed algorithm is ranked first in terms of the average accuracy (90.29%), while the CT-LS-SVM algorithm [14] came second (88.32%), the R-CSP-A [13] is third (83.9%) and so on. The SSFO [10] is the last (73.50%). The results indicate that the proposed method achieves by 1.97% to 16.79% improvements over all the eight existing algorithms for dataset IVa.

The results obtained in this paper demonstrate that the proposed method outperforms all eight other existing algorithms in terms of the average accuracy rate for dataset IVa of BCI Competition III.

IV. CONCLUSIONS

This paper presents a cross-correlation-based logistic regression algorithm in EEG signal analysis for the classification of MI tasks. In the adopted framework, the features of the MI EEG signals are computed by the cross-correlation method and the extracted features are employed as the inputs to the logistic regression model for the classification. The 10-fold cross-validation procedure is used to estimate the performance of the proposed algorithm in terms of specificity, sensitivity and accuracy. The experimental evaluation is performed on dataset IVa from BCI Competition III. The performance of the proposed approach is compared with most recently reported eight methods. The experimental results demonstrate that our present method is superior to the eight methods in the literature. Hence, it is obvious that the cross-correlation technique is suitable for representative feature extraction from the MI data and the logistic regression is an efficient classifier to distinguish the features of the MI data. In the future, we will extend the proposed cross-correlation-based logistic regression algorithm to the multiclass classification problems.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtschelle and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysio.*, vol. 113, no. 6, pp. 767-791, June 2002.

[2] A. Asadi Ghanbari, M .R. Nazari Kousarrizi, M. Teshnehlab and M. Aliyari, "An evolutionary artefact rejection method for brain computer interface using ICA," *Int. J. of Elect. & Comp. Scien.*, vol. 9, no. 9, pp. 461-466, 2009.

[3] G. Pfurtscheller, G. R. Muller-putz, A. Schlogl, B. Graimann, R. Scherer and R. Leeb, "15 years of BCI reaserch at Graz university of technology: current projects," *IEEE Trans. Neur. Syst. Rehab. Eng.,* vol. 14, no. 2, pp. 205-210, 2006.

[4] D. J. McFarland, W. A. Sarnacki and J. R. Wolpaw, "Brain-computer interface (BCI) operation: optimizing information transfer rates," *Biol. Psycho.*, vol. 63, pp. 237-251, 2003.

[5] M. Dalponte, F. Bovolo and L. Bruzzone, "Automatic selection of frequency and time intervals for classification of EEG signals," *Elect. Lett.*, vol. 43, no. 25, pp. 1406-1408, 2007.

[6] N. Sharma, V.M. Pomeroy, and J.C. Baron, "Motor imagery: a backdoor to the motor imagery system after stroke?," *Stroke.*, vol. 37, pp. 1941-1952, 2006.

[7] T. Kayikcioglu and O. Aydemir, "A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data," *Patt. Recog. Lett.*, vol. 31, pp. 1207-1215, 2010.

[8] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *J. Neural Eng.*, vol. 4, pp. R1-R13, 2007.

[9] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification,* 2nd ed., John Wiley & Sons, 2001.

[10] X, Yong, R.K. Ward and G.E. Birch, "Sparse spatial filter optimization for EEG channel reduction in brain-computer interface," *ICASSP 2008*, pp. 417-420.

[11] H. Lu, K.N. Plataniotis and A.N. Venetsanopoulos, "Regularized common spatial patterns with generic learning for EEG signal classification," *31st Annual International Conference of the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6, 2009*, pp. 6599-6602.

[12] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Trans. on Biom. Eng.*, in press, July 2010.

[13] H. Lu, H. L. Eng, C. Guan, K. N. Plataniotis, and A.N. Venetsanopoulos, "Regularized common spatial patterns with aggregation for EEG classification in small-sample setting," *IEEE Trans. on Biom. Eng.,* vol. 57, no. 12, pp. 2936-2945, December 2010.

[14] Siuly, Y. Li, and P. Wen, "Clustering technique-based least square support vector machine for EEG signal classification," *Comp. Meth. and Prog. in Biom.*, in press, 2010.

[15] T. A. L. Wren, K. P. do, S. A. Rethlefsen, and B. Healy, "Cross-correlation as a method for comparing dynamic electromyography signals during gait," *J. of Biomech.*,vol. 39, pp. 2714-2718, 2006.

[16] S. Chandaka, A. Chatterjee and S. Munshi, "Cross-correlation aided support vector machine classifier for classification of EEG signals," *Exp. Syst. with Appl.*, vol. 36, pp. 1329-1336, 2009.

[17] S. Dutta, A. Chatterjee and S. Munshi, "Correlation techniques and least square support vector machine combine for frequency domain based ECG beat classification," *Med. Eng. and Phys.*, in press, 2010.

[18] M. N. Islam, *An introduction to statistics and probability,* 3rd ed., Mullick & brothers, Dhaka New Market, Dhaka-1205, pp. 160-161.

[19] W. Caesarendra, A. Widodo and B.S. Yang, "Application of relevance vector machine and logistic regression for machine degradation assessment," *Mech. Syst. and Sign. Proce.*,vol. 24, pp. 1161-1171, 2010.

[20] D. W. Hosmer and S. Lemeshow, *Applied logistic regression*, Wiley, New York, 1989.

[21] A. Subasi and E. Ercelebi, "Classification of EEG signals using neural network and logistic regression," *Comp. Meth. and Prog. in Biom.*, vol. 78, pp. 87-99, 2005.

[22] BCI competition III, http://www.bbci.de/competition/iii

[23] B. Blankertz, K. R. Muller, D. J. Krusierski, G. schalk, J. R. wolpaw, A. Schlgl, G. Pfurtscheller, , and N. Birbaumer, "The BCI competition III: validating alternative approaches to actual BCI problems," *IEEE Trans. on Neur. Syst. and Rehab. Eng.*, vol. 14 , no. 2,pp. 153-159, 2006.

[24] Siuly, Y. Li and P. Wen, "EEG signal classification based on simple random sampling technique with least square support vector machines," *Int. J. of Biom. Eng. and Tech.*, in press, 2010.

[25] Siuly, Y. Li and P. Wen, "Analysis and classification of EEG signals using a hybrid clustering technique," *Proceedings of the 2010 IEEE/ICME International Conference on Complex Medical Engineering (CME2010)*, pp. 34–39.

[26] Siuly, Y. Li and P. Wen, "Classification of EEG signals using sampling techniques and least square support vector machines," *RSKT 2009, LNCS 5589*, pp. 375-382.