



# Multiple Kernel Based Region Importance Learning for Neural Classification of Gait States from EEG Signals

Yuhang Zhang<sup>1,2</sup>, Saurabh Prasad<sup>2</sup>, Atilla Kilicarslan<sup>1</sup> and Jose L. Contreras-Vidal<sup>1\*</sup>

<sup>1</sup> Noninvasive Brain-Machine Interface Systems Lab, Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA, <sup>2</sup> Hyperspectral Image Analysis Lab, Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA

## OPEN ACCESS

### Edited by:

Ioan Opris,  
University of Miami School of  
Medicine, USA

### Reviewed by:

Damien Coyle,  
Ulster University, UK  
Lei Ding,  
University of Oklahoma, USA

### \*Correspondence:

Jose L. Contreras-Vidal  
jlcontreras-vidal@uh.edu

### Specialty section:

This article was submitted to  
Neural Technology,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 01 December 2016

**Accepted:** 15 March 2017

**Published:** 03 April 2017

### Citation:

Zhang Y, Prasad S, Kilicarslan A and  
Contreras-Vidal JL (2017) Multiple  
Kernel Based Region Importance  
Learning for Neural Classification of  
Gait States from EEG Signals.  
*Front. Neurosci.* 11:170.  
doi: 10.3389/fnins.2017.00170

With the development of Brain Machine Interface (BMI) systems, people with motor disabilities are able to control external devices to help them restore movement abilities. Longitudinal validation of these systems is critical not only to assess long-term performance reliability but also to investigate adaptations in electrocortical patterns due to learning to use the BMI system. In this paper, we decode the patterns of user's intended gait states (e.g., stop, walk, turn left, and turn right) from scalp electroencephalography (EEG) signals and simultaneously learn the relative importance of different brain areas by using the multiple kernel learning (MKL) algorithm. The region of importance (ROI) is identified during training the MKL for classification. The efficacy of the proposed method is validated by classifying different movement intentions from two subjects—an able-bodied and a spinal cord injury (SCI) subject. The preliminary results demonstrate that frontal and fronto-central regions are the most important regions for the tested subjects performing gait movements, which is consistent with the brain regions hypothesized to be involved in the control of lower-limb movements. However, we observed some regional changes comparing the able-bodied and the SCI subject. Moreover, in the longitudinal experiments, our findings exhibit the cortical plasticity triggered by the BMI use, as the classification accuracy and the weights for important regions—in sensor space—generally increased, as the user learned to control the exoskeleton for movement over multiple sessions.

**Keywords:** brain machine interface (BMI), neural classification, electroencephalography (EEG), machine learning, multiple kernel learning

## 1. INTRODUCTION

Brain Machine Interface (BMI) systems have attracted extensive attention in the past decade, because of their potential in improving human life, especially for those who are affected by motor disabilities. Since gait deficits are commonly associated with spinal cord injuries (SCI), limb loss, and neurodegenerative diseases, there is a need to investigate innovative therapies to restore gait in such patients. Exoskeletons have become prominent tools for the rehabilitation of SCI and stroke patients (Sale et al., 2012; Venkatakrishnan et al., 2014). BMIs have been deployed to infer the

user's intent from his/her brain activity to generate output signals to control powered exoskeletons for upper and lower limb rehabilitation (Noda et al., 2012; Contreras-Vidal and Grossman, 2013; Kilicarslan et al., 2013; French, 2014; Venkatakrishnan et al., 2014). In Presacco et al. (2011), Presacco et al. showed decoding of gait kinematics during treadmill walking from EEG of able-bodied subjects with accuracies comparable to that from a similar study in non-human primates with electrodes implanted in their brains (Fitzsimmons et al., 2009). Further, in Kilicarslan et al. (2013), a paraplegic subject's motion intentions were accurately decoded using Locality Fisher Discriminant Analysis and a Gaussian Mixture Model (LFDA-GMM) for two different gait tasks (i.e., repeated walking-turning right-turning left motions and sit-rest-stand motions). The model enabled the closed-loop EEG-based BMI system to control a robotic exoskeleton (NeuroREX) in real-time, resulting in independent walking for the paraplegic user.

To control a device via BMI, different brain activity patterns produced by a user need to be accurately identified by a neural interface system and translated into appropriate commands (Contreras-Vidal et al., 2015). Discrete decoding (neural classification) of intent from EEG signals can be considered as a pattern recognition problem, and advanced machine learning techniques are needed to accurately translate the brain electrical activities to meaningful control commands. Many machine learning methods [e.g., linear discriminant analysis (LDA), support vector machine (SVM), Bayesian classifiers] have been applied for classifying EEG signals in different BMI applications (Kilicarslan et al., 2013; Niazi et al., 2013; Leamy et al., 2014; Lew et al., 2014; Hortal et al., 2015; Jiang et al., 2015). However, most of them serve as a "black box" in that we do not know how the brain activity changes during long-term BMI use nor how the brain regions contribute to the classification process while people perform different tasks. The human brain consists of over 100 billion cells, typically divided into regions by neuroanatomists. Different regions have their specific functionalities while coordinating together to accomplish everyday tasks. Moreover, the specific contributions of brain regions to classification may change due to learning a BMI. Therefore, it is important to identify and track these changes to increase our understanding of brain function, BMI learning and performance. In that context, the hypothesis of this research is that different brain regions contribute differentially to BMI learning and control of robot assisted lower-limb movements—we are interested in learning the importance of these regions for neural classification of gait states.

Kernel learning methods have been effectively applied for many machine learning problems, including feature selection, data regression and classification for EEG signals (Garrett et al., 2003; Lal et al., 2004; Lotte et al., 2007). SVM is one of the most popular kernel methods for pattern recognition. However, a problem with using the standard SVM in BMI applications is that it provides no insight about the importance of distinct features, and thus has little knowledge about the biophysical properties of relevant features used in decoding/classification. Multiple kernel learning (MKL), which makes use of a combination of basis kernels to represent different types of features or data, have

been shown to outperform traditional single-kernel machines in different aspects (Sonnenburg et al., 2006; Tian et al., 2012; Samek et al., 2013; Li et al., 2014). The main advantage of using MKL over SVM is that MKL can simultaneously learn the classifier and the optimal weights for basis kernels. In this paper, we investigate and make use of this property to simultaneously decode gait states from multi-channel EEG signals and learn the relative importance of different scalp brain areas. Particularly, we build a composite kernel based on a linear combination of basis kernels, in which each basis kernel can be represented by a group of electrodes corresponding to selected regions of interest (ROIs), and consequently contribute unique biophysical information.

The primary goal of this research is to show the feasibility of simultaneously classifying the pattern of user's internal gait states (e.g., stop, walk, turn left, turn right) from the EEG signals and learning the relative importance of different scalp brain areas. Previous studies have shown that low delta band (0.1–2 Hz) EEG contains intended movement-related information for decoding the kinematics of lower limb or gait states (Presacco et al., 2011, 2012; Jorquera et al., 2013; Kilicarslan et al., 2013; Bulea et al., 2014; Luu et al., 2016). For example, in Presacco et al. (2011), Presacco et al. (2012), and Luu et al. (2016), it was shown that delta band EEG contains information about gait movement kinematics that can be decoded using Wiener or Kalman filters. In Kilicarslan et al. (2013), Jorquera et al. (2013), and Bulea et al. (2014), it was shown that movement-type (e.g., "stop," "go," etc.) classifiers can be designed based on delta band EEG signals. Another study (Velu and de Sa, 2013) showed that features corresponding to frequencies less than 2 Hz were the most heavily weighted during single trial classification of walking and pointing direction. Inspired by the above findings, in this study, we utilize delta band (0.1–2 Hz) EEG to build our basis kernels (feature matrices) for neural classification of gait states.

The other goal of the research is to compare the brain regions employed for classifying movement intents from able-bodied subjects and individuals with spinal cord injury (SCI) given differences in neural activity across these populations. Studies have shown that SCI can cause widespread and sustained brain inflammation that leads to progressive loss of brain cells in key brain regions with associated cognitive problems (Wu et al., 2014a,b). Cramer et al. have found that in patients with complete SCI, many features of normal motor system function are preserved, however, the volume and patterns of activation and the modulation of function with change in task are abnormal and absent, respectively, in patients with SCI (Cramer et al., 2005). In this preliminary study, we collected EEG data from a SCI volunteer over multiple sessions to compare the classification results with an able-bodied subject on the important brain regions during learning.

The remainder of the paper is organized as follows. Section 2 introduces our methodology for region importance learning, including experimental protocol, data acquisition, processing and analysis. In particular, we introduce the MKL algorithm and how we apply it to learn the importance of different brain regions. In section 3, we validate the efficacy of the proposed work via experiments using four-class single session and two-class longitudinal EEG data.

Section 4 presents our discussions from analyzing the experimental results. Finally, concluding remarks are provided in Section 5.

## 2. METHODOLOGY

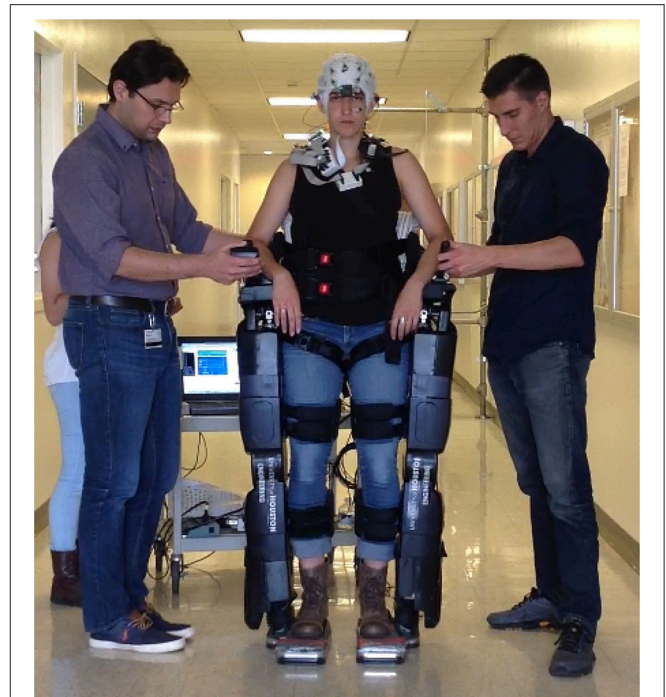
### 2.1. Experimental Protocols and Tasks

The experimental protocols were approved by the Institutional Review Board of the University of Houston. After giving written informed consent, an able-bodied subject and an SCI subject (both male) were fitted with a wearable powered exoskeleton (REX, REX Bionics Ltd, New Zealand) and an EEG-based BMI (Kilicarslan et al., 2013). For data collection, users were asked to perform motor imagery of locomotive movements while following and completing a path marked on the ground with the robot controlled by an operator remotely. This allowed synchronized motion and EEG data while securing user engagement. There were two tasks in this research. Task 1 was a four-class, single session task in which the subjects performed different movements, i.e., walking forward, turning right, turning left and stop, following the marked path on the ground. In Task 2, subjects only executed walking and stop motions according to audible beep instructions. Each trial contained at least 10 stop-to-walk or walk-to-stop transitions. The subjects were trained over multiple sessions in a 30 days period to control the exoskeleton to perform these motions.

### 2.2. Data Acquisition and Processing

Multichannel active-electrode EEG (64 channels) was recorded by combining two 32-channel amplifiers (actiCap system, Brain Products GmbH, Germany). The electrodes were placed and labeled in accordance with the extended 10–20 international system. A wireless interface (MOVE system, Brain Products GmbH, Germany) was used to transmit data (sampled at 100 Hz) to the host PC. **Figure 1** shows a volunteer controlling NeuroRex via the EEG BMI system.

We took a careful approach in regard to potential motion artifacts aiding decoding. First, we used good engineering measurement practices (Nathan and Contreras-Vidal, 2015), including EEG cap set-up and medical-grade mesh to fixate individual electrode wirings that can induce motion artifacts; second, we deployed a wireless active-electrode EEG system to increase the signal to noise ratio (signals are amplified directly at the electrode location) and help mitigate motion artifacts; third, we have shown that the delta band EEG contains negligible motion artifacts at the gait speeds tested in the study (Nathan and Contreras-Vidal, 2015); fourth, we applied the Artifact Subspace Reduction (ASR, an automated artifact rejection method Mullen et al., 2013; Bulea et al., 2014) and compared classification accuracies with and without ASR, to assess the potential effects of motion artifacts but did not find significant changes on classification accuracies suggesting that motion artifacts, if any, did not affect decoding. The acquired data were then filtered in the 0.1–2 Hz range using a second order Butterworth filter and standardized (z-score) in a data preprocessing step.



**FIGURE 1 |** A volunteer controlling NeuroRex via the EEG BMI system.

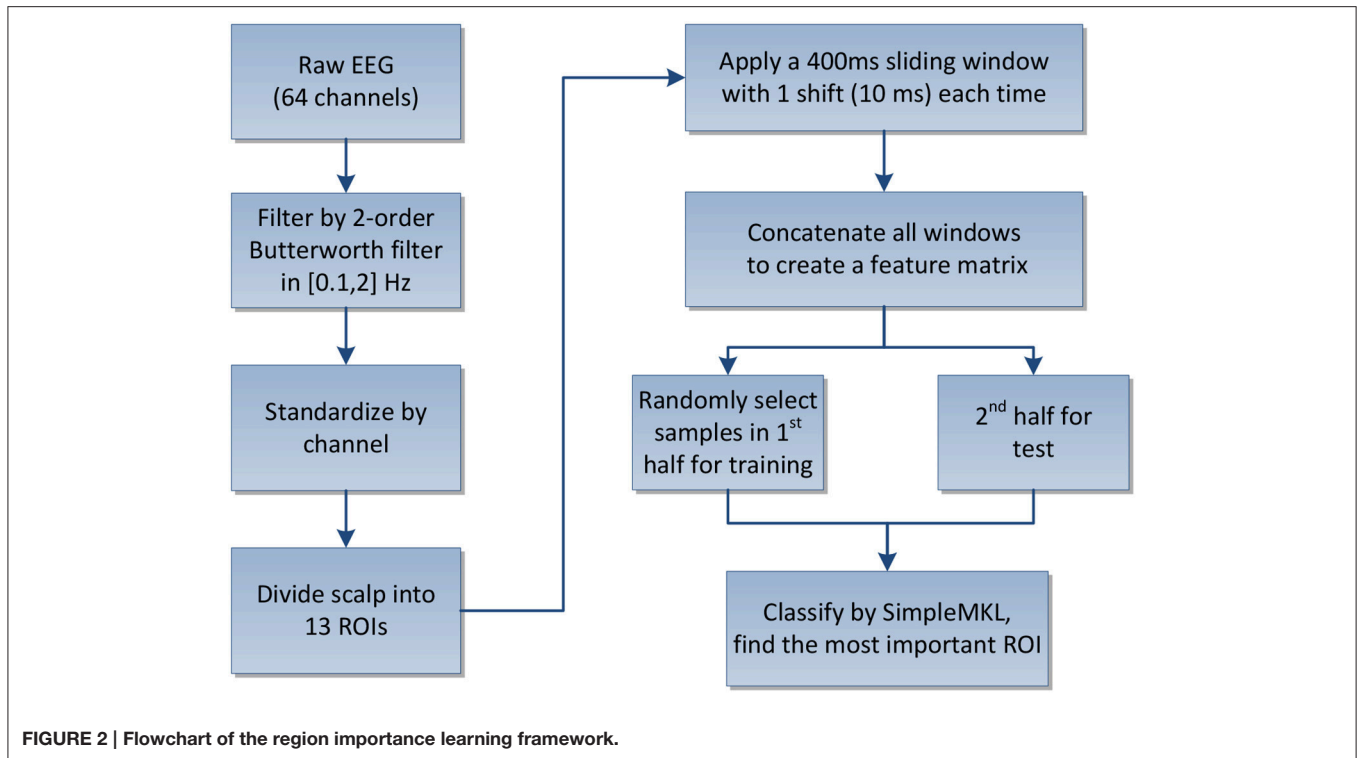
### 2.3. Region Importance Learning Framework

We conducted separate experiments of the above two tasks to interpret the use of kernel weights in MKL as an indicator of the region importance in classification of user's movement intention from EEG signals. After the signals were pre-processed, 64 channels were divided into 13 ROIs as described in Section 2.4.1. The features were then extracted by applying a 400 ms sliding window on each channel with 1 shift (10 ms) each time to acquire the amplitude modulations and concatenated as a feature matrix. To better meet the data process in real world, we divided the labeled samples into two halves for supervised learning. We randomly select 500 samples from the first half of the labeled samples for training, and the remaining half were used for testing and evaluation. The testing process was repeated 10 times and the metric for evaluating the classification results is the average overall accuracy (OA). The flowchart of the proposed framework is shown in **Figure 2**.

### 2.4. Learning and Classification Methods

#### 2.4.1. Brain Scalp Regions

The analysis and interpretation of EEG measurements depend upon the correspondence of electrode scalp coordinates to structural and functional regions of the brain (Giacometti et al., 2014; Gentili et al., 2015). For example, Giacometti et al. (2014) showed that EEG electrode proximity maps intersect with EEG sensitivity maps of the human brain, allowing the use of proximity maps to inform the cortical origin of scalp recordings. Furthermore, intersection of structural and functional regions of the brain with cortical proximity parcellations can be used to



show the correspondences between scalp electrode coordinates and potential regions of interest in the human cortex (Giacometti et al., 2014).

In this research, we investigate the importance of brain areas for the lower-limb movement neural classification task for both able-bodied and SCI subjects. Specifically, the brain scalp is divided into 13 topographical regions of interest (ROIs) (Kranzioch et al., 2008; Gobel et al., 2011), which are anterior frontal (AF), left fronto-central (LFC), midline fronto-central (MFC), right fronto-central (RFC), left centro-parietal (LCP), midline centro-parietal (MCP), right centro-parietal (RCP), left parieto-occipital (LPO), middle parieto-occipital (MPO), right parieto-occipital (RPO), left temporal (LT), right temporal (RT) and Occipital (O). **Figure 3** and **Table 1** show the partition of the scalp and the name for each ROI.

#### 2.4.2. Kernel-Based Learning Methods Foundation

Kernel-based learning methods have been widely applied for various machine learning tasks. The reason for its popularity is that it easily extends the linear classifier to nonlinear decision surfaces using the “kernel trick.” All the kernel methods make use of the “kernel trick” to map the data  $X = \{x_1, x_2, \dots, x_N\}$  from the input space to a higher dimensional feature space  $\mathcal{H}$  (i.e., Reproducing Kernel Hilbert Space (RKHS)) as  $\Phi: \mathbb{R}^d \rightarrow \mathcal{H}, x \rightarrow \Phi(x)$ , so that the original non-linear data are linear separable in such feature space. The kernel mapping is defined as:

$$K(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle, \quad (1)$$

where  $\langle \cdot, \cdot \rangle$  is the inner product of two vectors.

SVM is one of the most popular kernel-based classifier (Vapnik and Vapnik, 1998). The underlying principle of SVM is to simultaneously minimize the empirical classification error and maximize the geometric margin of the linear separation surface. The optimization problem for SVM classification is formulated as:

$$\begin{aligned} \min_{w, \xi_i, b} J(w, \xi_i, b) &= \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ \text{s.t. } &\begin{cases} y_i (\langle w, \Phi(x_i) \rangle + b) \geq 1 - \xi_i \\ \xi_i \geq 0, \forall i = 1, 2, \dots, N \end{cases} \end{aligned} \quad (2)$$

where  $C$  is a constant which controls the balance between the margin and empirical loss,  $\xi_i$  are slack variables which measure the degree of misclassification, and  $\|w\|^2$  is inversely related to the margin to the hyperplane.

In most kernel-based learning methods, performance is greatly affected by the choice of kernel function and related kernel hyper-parameters. The standard SVM only utilizes a single kernel function with fixed parameters, which necessitates model selection for good classification performance. Besides, using a fixed kernel may be suboptimal, since different sources of data may have different representations of the phenomena of interest, and hence the similarity should not be measured via the same kernel function.

#### 2.4.3. Multiple Kernel Learning (MKL)

In recent works, MKL has been shown to outperform traditional single-kernel SVMs in many cases, especially for classification and feature fusion problems (Sonnenburg et al., 2006; Tian et al., 2012; Samek et al., 2013; Li et al., 2014; Zhang et al.,



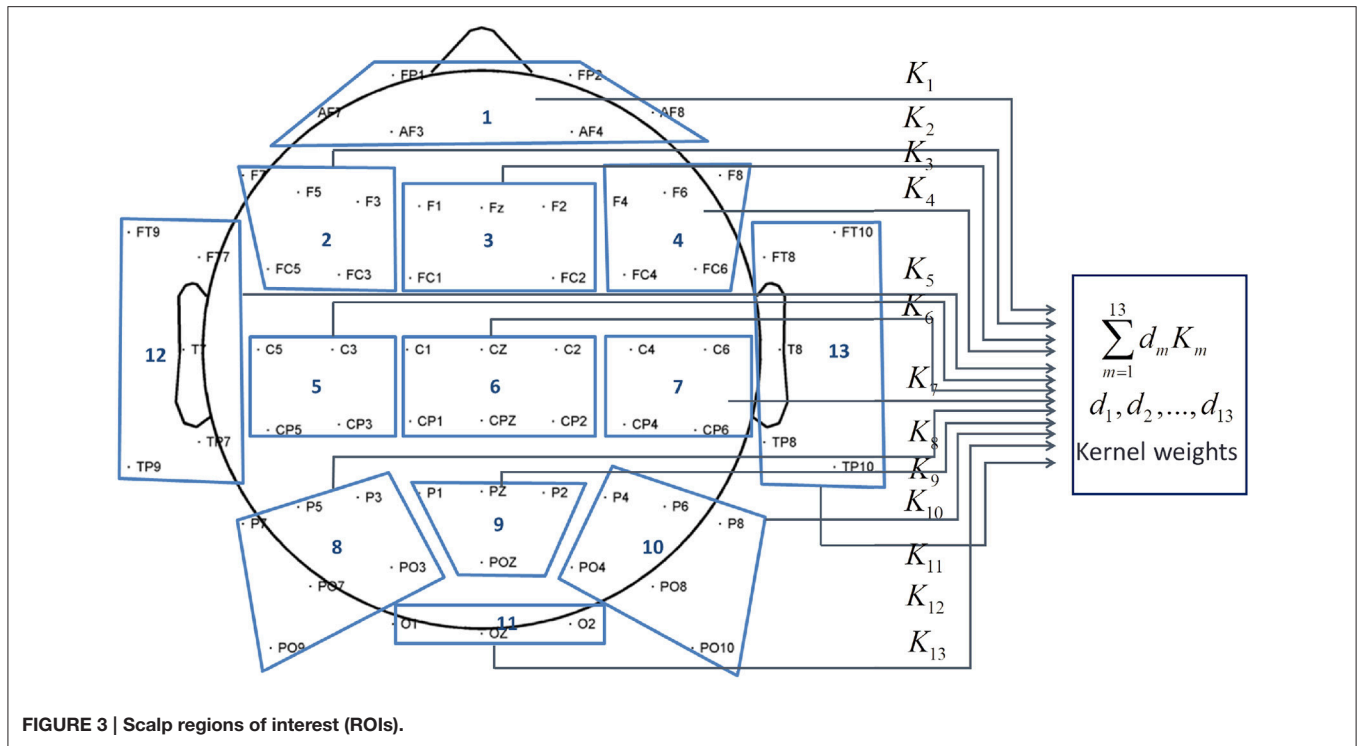


TABLE 1 | Scalp ROI names.

Index	ROI name	Index	ROI name
1	Anterior Frontal (AF)	8	Left Parieto-Occipital (LPO)
2	Left Fronto-Central (LFC)	9	Middle Parieto-Occipital (MPO)
3	Midline Fronto-Central (MFC)	10	Right Parieto-Occipital (RPO)
4	Right Fronto-Central (RFC)	11	Left Temporal (LT)
5	Left Centro-Parietal (LCP)	12	Right Temporal (RT)
6	Midline Centro-Parietal (MCP)	13	Occipital (O)
7	Right Centro-Parietal (RCP)		

2015). In this paper, we employ MKL to infer information about electrode relevance by observing the kernel weights learned from training the machine for classification. Each “group” of features is assigned a basis kernel, and the linear combination of all basis kernels is optimized through gradient descent on the SVM objective function. The optimization of multiple kernels works as a feature selector providing a weighted ranking of the importance of its components.

We consider the above 13 ROIs as generating a 13-source input. For a specific source  $p$ , the combined kernel function  $K$  between two samples  $x_i^p$  and  $x_j^p$  can be represented as

$$K(x_i^p, x_j^p) = \sum_{m=1}^M d_m K_m(x_i^p, x_j^p) \quad (3)$$

s.t.  $d_m \geq 0$ , and  $\sum_{m=1}^M d_m = 1$ ,

where  $M$  is the number of candidate basis kernels representing different kernel parameters,  $K_m$  is the  $m$ -th basis kernel and  $d_m$  is the weight for it. Weights can be estimated through cross-validation, which is computationally demanding when the number of basis kernels (i.e., feature sets or data sources) is large. An alternative strategy, which we adopt in this work, is based on the SimpleMKL algorithm (Rakotomamonjy et al., 2008). It optimizes the weights automatically in a learning problem by utilizing the gradient descent approach. Based on the SVM optimization problem, the SimpleMKL learning problem is expressed as:

$$\min_d J(d), \text{ s.t. } d_m \geq 0, \text{ and } \sum_{m=1}^M d_m = 1$$

$$J(d) = \begin{cases} \min_{w,b,\xi} \frac{1}{2} \sum_{m=1}^M \frac{1}{d_m} \|w_m\|^2 + C \sum_{i=1}^N \xi_i \\ \text{s.t. } y_i \left( \sum_{m=1}^M \langle w_m, \Phi_m(x_i^p) \rangle + b \right) \geq 1 - \xi_i \\ \xi_i \geq 0, \forall i = 1, 2, \dots, N, \end{cases} \quad (4)$$

where  $\Phi_m(x_i^p)$  is the kernel mapping function of  $x_i^p$ ,  $w_m$  is the weight vector of the  $m$ <sup>th</sup> decision hyperplane,  $C$  is the regularization parameter controlling the generalization capabilities of the classifier, and  $\xi_i$  is a positive slack variable.

The objective function is a constrained optimization problem, which can be transformed into a dual form  $L(\alpha_i, \alpha_j)$  using Lagrange multipliers  $\alpha_i, \alpha_j$ . Then the kernel weight  $d_m$  can be optimized by updating it along the gradient descent direction of  $L(\alpha_i, \alpha_j)$  as  $d \leftarrow d + \gamma D$ , where  $\gamma$  is the step length,  $D$  is the descent direction of  $L(\alpha_i, \alpha_j)$ , and  $d = [d_1, d_2, \dots, d_M]^T$  is the

kernel weight vector. Following this optimization procedure and after several iterations, SimpleMKL provides the optimal kernel weight for each basis kernel that indicates the importance of a particular brain region in classification of gait states.

## 2.5. Parameter Settings

In the experiments, RBF kernels defined as  $K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2})$  were used with relative width parameter  $\sigma$ . In the multiple kernel setting, we did not select a specific kernel parameter; instead, we defined a set of different values as candidate input parameters. We can build several basis kernels with different values of  $\sigma$  for each source of input, however, the number of parameters should be kept small to reduce the computational complexity and memory requirements. In particular, four basis kernels with  $\sigma = [0.1, 0.5, 1, 1.5]$  were considered for all sources. This range of values was found to be reasonable after applying kernel alignment (Shawe-Taylor and Kandola, 2002) using cross-validation. The penalty parameter was then selected by cross-validation in the range of  $[2^{-1}, \dots, 2^{15}]$ . For further information of MKL experimental settings, we refer readers to Zhang et al. (2015). All the experiments were implemented in Matlab R2014a using the SimpleMKL toolbox (Rakotomamonjy, 2008).

## 3. RESULTS

### 3.1. Four-Class, Single Session Classification Results

First, we compare the kernel weights optimized by SimpleMKL algorithm for defined ROIs from the able-bodied subject and the SCI subject in a four-class task. Four motion classes for classification are walking forward, turning left, turning right and stop. The boxplots and topoplots of optimized kernel weights for different ROIs are shown in **Figure 4**. The average classification accuracies were 74.5% and 68.4% for the able-bodied and the SCI subject, respectively.

From the results, it is observed that the fronto-central scalp regions (MFC, RFC) have the highest weights among all ROIs, which included scalp areas associated with the motor planning and the lower-limb neural representation (Leeb et al., 2013). Interestingly, for the able-bodied subject, the MFC ROI showed the highest relevance to gait decoding with RFC being the closest area in importance. In contrast, in the case of the SCI subject, the order of importance was reversed, with RFC showing the highest relevance followed by MFC. We also note that for the SCI subject, ROIs LFC, LCP, MCP, and RCP also showed relatively higher weights than for the healthy control subject, while the remaining ROIs have low weights for both subjects. Clearly, the cortical representation for the gait movements was more compact and strong for the able-bodied subject than the SCI user. These results demonstrate that MKL can be efficiently used to infer the importance of different groups of features and thus suggest different roles in the representation of gait for different scalp brain areas.

Further, we give some insights of the class-wise results regarding different movement intentions. We show the confusion

matrices in terms of class-wise accuracies and misclassification rates in **Figure 5**. Generally, the stop intention is the most difficult to decode—it was misclassified as walking forward in many situations. Turning right always has a high accuracy for both able-bodied and SCI subjects compared to the other classes. We note that all class-wise accuracies are above channel level—which is 25% for this problem.

### 3.2. Two-Class, Multiple Sessions Classification Results

Second, we conducted a longitudinal experiment for the two-class (i.e., walk and stop) classification problem. We quantified electrode relevance changes across sessions to examine neural signatures that may indicate the cortical plasticity triggered by BMI use. We first plot the weight changes along 9 sessions over a period of 30 days for the able-bodied (a different subject as in task 1) and the SCI subjects (the same subject as in task 1) in **Figures 6, 7**, respectively. As depicted in the scalp maps, the weights change dramatically in the first several daily sessions, while becoming more stable in the later sessions. Similar to the previous results, the frontal scalp regions get the highest weights among all ROIs after training the user to control the exoskeleton for several sessions. Specially, for the SCI subject, RFC (ROI 4) has the highest weight, while LCP (ROI 5) also has relative high weight. For the able-bodied subject, the final important region is determined as MFC (ROI 3). Thus, the SCI subject used different brain regions to operate the BMI system when compared with the able-bodied user.

Since ROIs 4, 5 and ROI 3 were determined as the most significant regions for the SCI subject and the able-bodied subject, respectively, for classification of gait states, we further evaluated the overall accuracy and kernel weight for these ROIs as a function of session. The linear fit of the relations between overall accuracy (or weights for the selected ROI) and daily BMI sessions are shown in **Figures 8, 9** for the SCI subject and the able-bodied subject, respectively. From the results, we can see the classification accuracy generally increases as a function of session. At the same time, the weight for ROI 4, 5 (ROI 3) also showed a trend of increasing along the session over a period of 30 days. We calculated the  $R^2$  and  $p$ -value as an indicator of how well the data fitted to the regression line and the significance of the results to the hypothesis, respectively.

## 4. DISCUSSION

Previous study (Kilicarslan et al., 2013) has shown the feasibility of classifying user movement intentions using EEG signals in the delta band (0.1–2 Hz) based on a GMM classifier and achieved a high offline evaluation accuracies for the 3-class tasks. In this study, we extend the study of intended motions from three classes to four classes and set the research as a longitudinal study. The experimental results demonstrate that by properly weighting the importance of the features, MKL can be used as an efficient decoder to predict user's movement intentions. For the tested subjects, the overall accuracies reached above 90% for the two-class classification task and above 65% for the more complicated

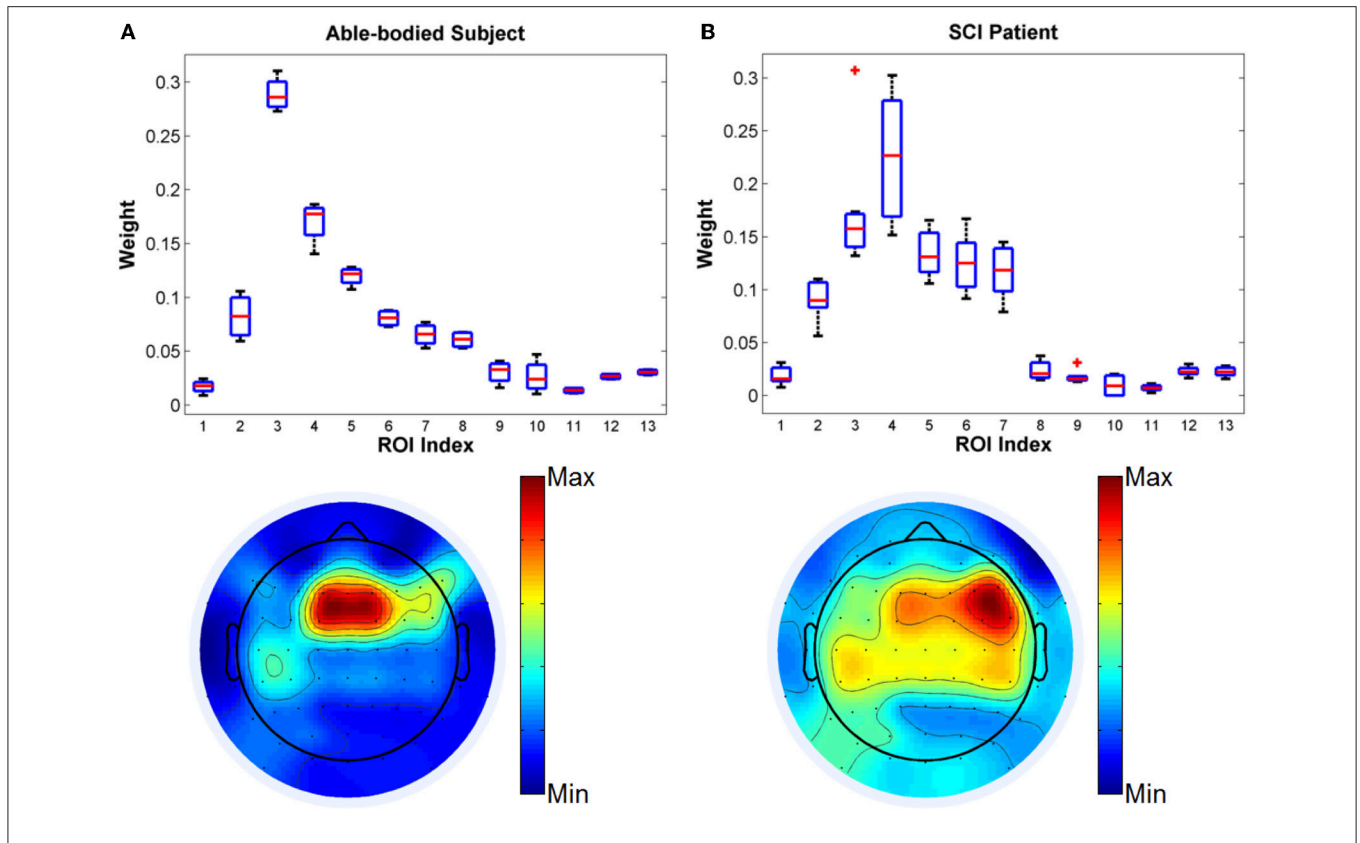


FIGURE 4 | Comparison of kernel weights for different ROIs from (A) able-bodied subject and (B) SCI subject in Task 1.

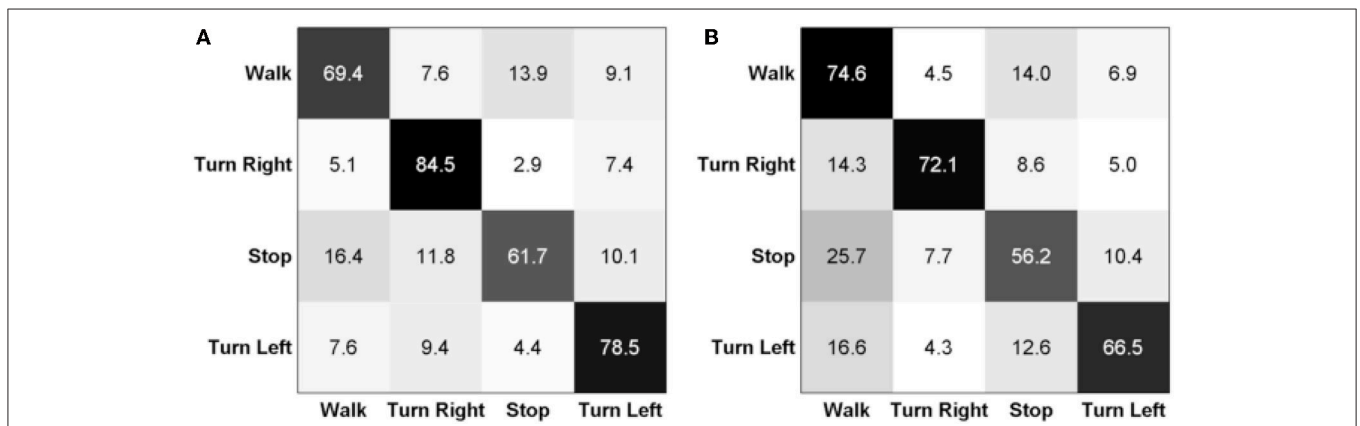
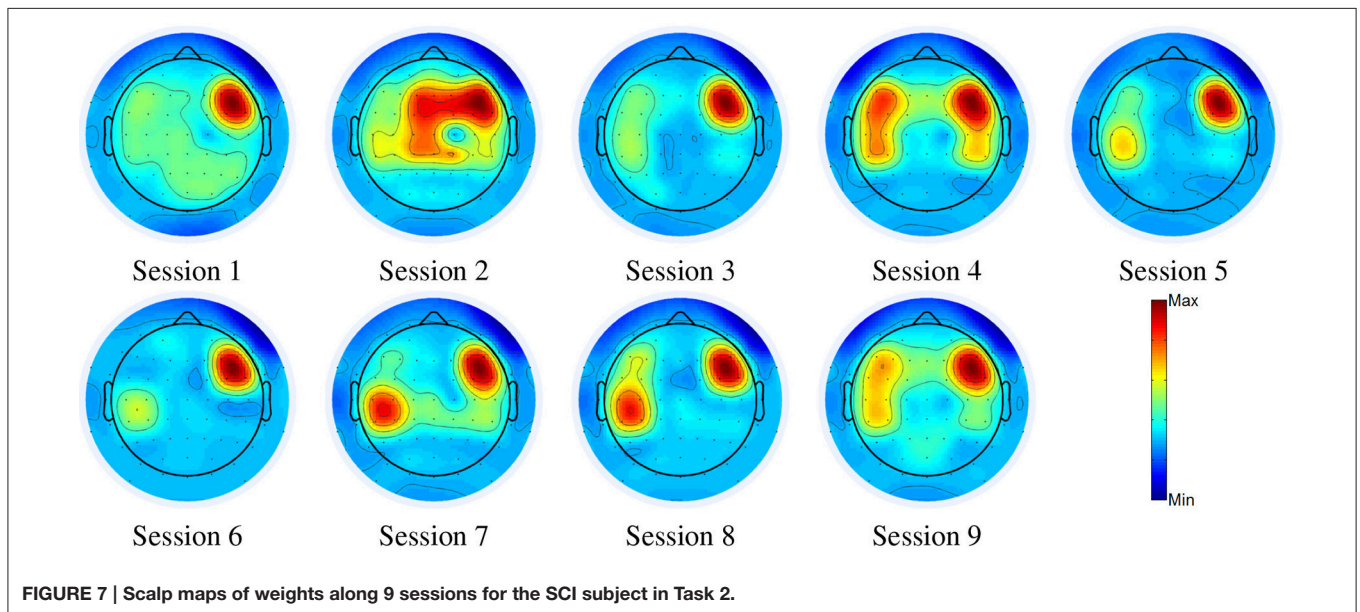
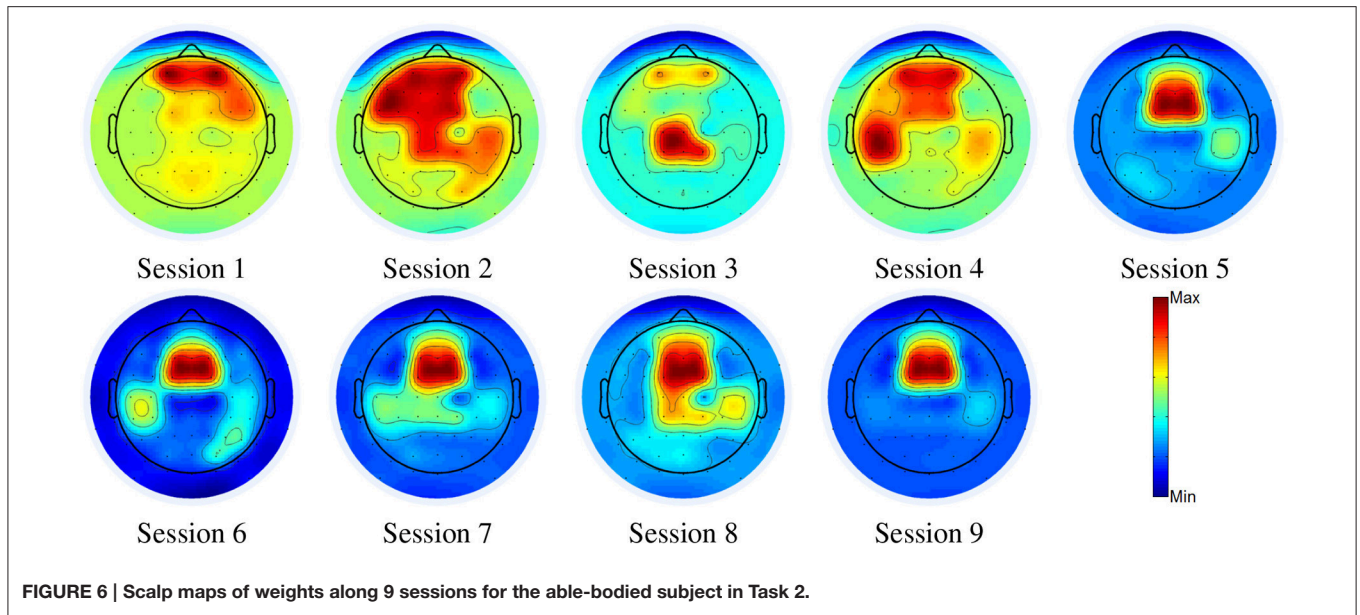


FIGURE 5 | Confusion matrices (%) for (A) able-bodied subject and (B) SCI subject in Task 1.

four-class classification task. Compared to some commonly used machine learning approaches (i.e., Bayesian classifiers, LDA, SVM) in BMI, MKL has the following advantages (1) Unlike the LDA and Bayesian classifiers, MKL does not need to make assumptions on the data distribution. MKL is a member of kernel learning methods, which utilizes a linear combination of kernels and transforms the original data into an appropriate (kernel) feature space. Thus, all beneficial properties (e.g., optimality) of linear classifiers are maintained, while MKL is also efficient when

the data are non-linear in the input space. (2) MKL is a robust learning method in the high dimensional space (Bach, 2008). In Kilicarslan et al. (2013), a sliding window was used on all 64 channels to extract features, it resulted in a 1,280 dimensional space, and a dimensionality reduction technique was required to reduce the dimension of the data decoding by GMM. Similar to the feature extraction step in Kilicarslan et al. (2013), we applied a 400 ms sliding window to extract the EEG delta band amplitude as input to the classifier. Differently, we divided the

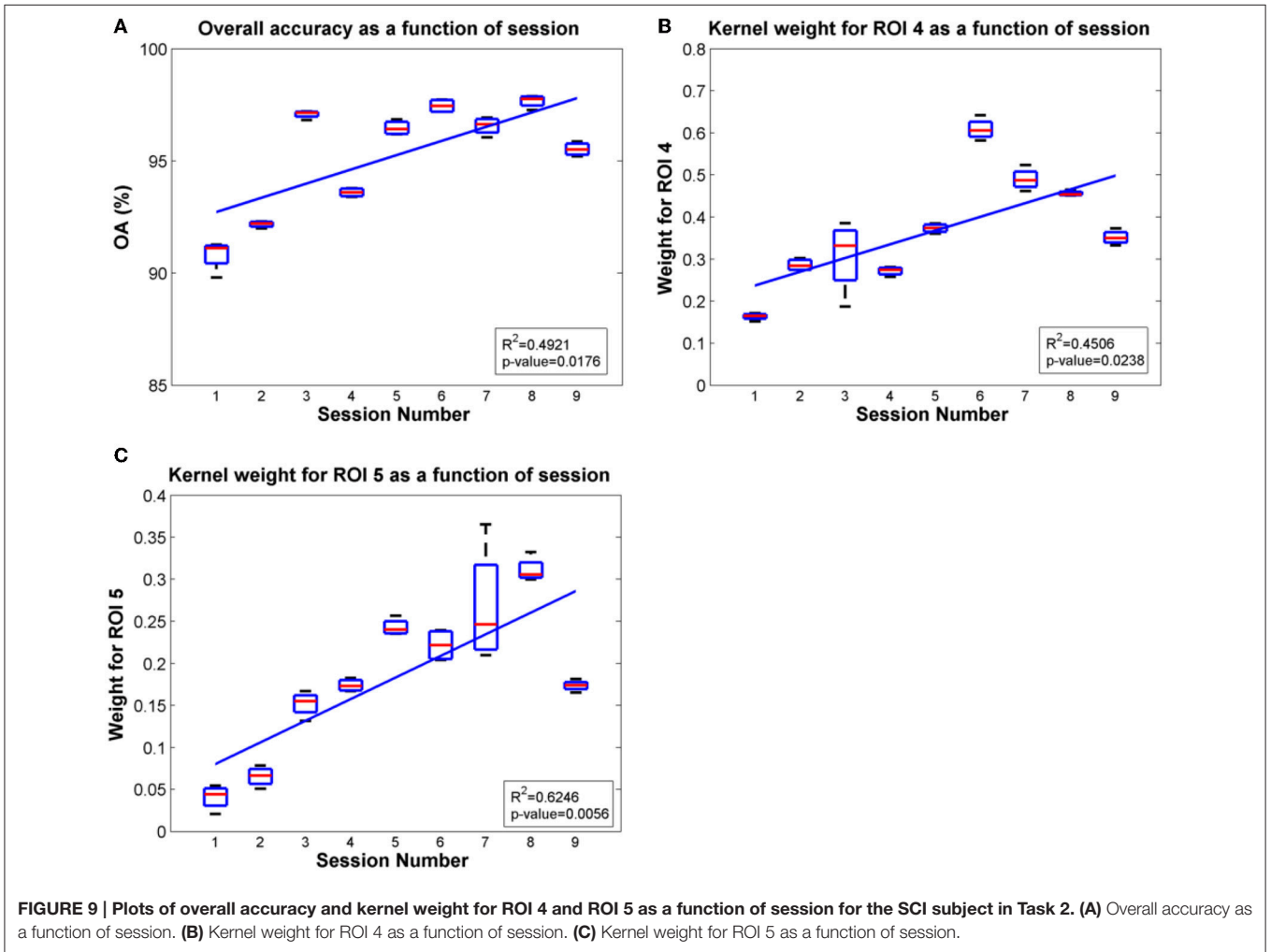
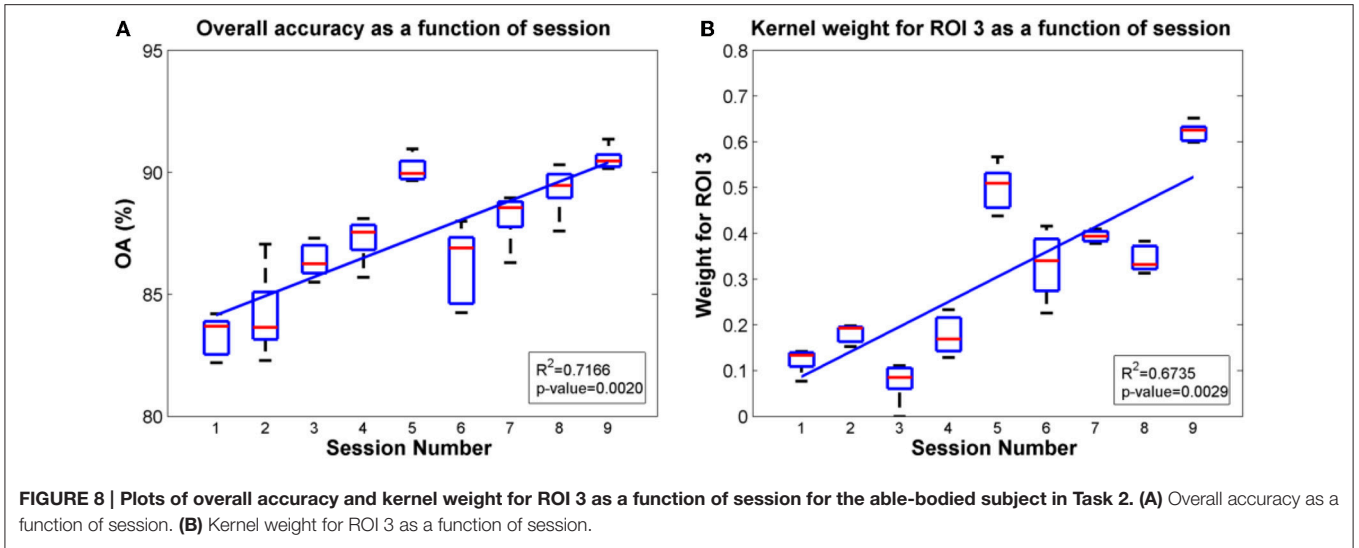


channels into different groups and extract features from each group. Thus, the resulting dimension for each group of data is at most 240, which is much lower than 1,280, and a dimensionality reduction method is not necessary for classification by MKL. (3) MKL can also be used to infer the importance of different groups of features, which is not feasible in other machine learning methods. The weight for each group is initialized uniformly at the beginning and optimized during the gradient descent in the MKL algorithm. MKL ranks sets of features corresponding to the meaningful features for solving the classification problem, and the results indicate importance in the representation of movement for different scalp brain areas.

Comparing the results from the SCI subject and the able-bodied subject, we observe the most important brain region

changing from the midline fronto-central to the right fronto-central in both tasks. This could be due to the loss of brain cells and degraded cerebral cortex dynamics or lack of afferent input after spinal cord injury. For example, changes in movement-related cortical potentials have been noted after SCI and correlated with the severity of the injury (Boord et al., 2008; Gourab and Schmit, 2010). Moreover, altered spontaneous neuronal activity following SCI has been characterized by a shift in the dominant spectral power peak toward lower frequencies, including in primary and secondary somatosensory cortices (Tran et al., 2004; Sarnthein et al., 2006). Noted that the number of subjects participated in the study was limited, there might be individual variations on the classification results. However, as the classification results were based on standard cross-validation





procedure, we believe the proposed approach and model can be generalized to other subjects.

In the longitudinal experiments, we found that the subjects were adapting to the BMI gait task in the first several sessions, so that the brain regions used for neural classification were not stable, which was reflected in the moderate classification performance. After several sessions of training, as the subjects learned to control the exoskeleton for movement, we observed the channels employed for movement classification converged to specific regions—the midline fronto-central areas for the able-bodied subject and the right fronto-central/left centro-parietal areas for the SCI user. In addition, the classification accuracy generally increased along session, and interestingly the weights for the important regions also increased. This demonstrates the cortical plasticity triggered by the BMI use, as the user gradually learns to control the exoskeleton for movement.

## 5. CONCLUSION

In this paper, we presented the feasibility of simultaneously classifying the pattern of user's internal gait states from the EEG signals and learning the relative importance of different scalp brain areas based on the MKL algorithm. The MKL has the advantages of learning the classifier and the optimal kernel weights simultaneously. We investigated these properties and applied the MKL classifier to infer the relative importance of different groups of features (different sources of information) in a BMI application to classify one's motion intention from the EEG signals. The experimental

results demonstrated that the frontal/fronto-central regions were the most important regions for classifying gait states of the tested subjects, which is consistent with the brain regions hypothesized to be involved in the control of lower-limb movements. By comparing the results from the SCI subject and the able-bodied subject, the important regions were observed to change, which could be due to the loss of brain cells and degraded cerebral cortex dynamics or lack of afferent input after spinal cord injury. In the longitudinal experiment, while the user learned to control the exoskeleton for movement over multiple sessions, the classification accuracy increased and the weights for important regions stabilized. These findings suggest the cortical plasticity triggered by the BMI use, which will be investigated further in the future study.

## AUTHOR CONTRIBUTIONS

YZ developed the learning method, processed and analyzed the data, and wrote the manuscript. SP supervised development of work, helped in data interpretation and manuscript edit. AK helped to acquire and interpret the data. JC supervised development of work, helped in data interpretation and manuscript edit and evaluation.

## FUNDING

This work was supported in part by the NIH Award R01 NS075889, Mission Connect—A TIRR Foundation.

## REFERENCES

- Bach, F. R. (2008). Consistency of the group lasso and multiple kernel learning. *J. Mach. Learn. Res.* 9, 1179–1225.
- Boord, P., Siddall, P., Tran, Y., Herbert, D., Middleton, J., and Craig, A. (2008). Electroencephalographic slowing and reduced reactivity in neuropathic pain following spinal cord injury. *Spinal Cord* 46, 118–123. doi: 10.1038/sj.sc.3102077
- Bulea, T. C., Prasad, S., Kilicarslan, A., and Contreras-Vidal, J. L. (2014). Sitting and standing intention can be decoded from scalp eeg recorded prior to movement execution. *Front. Neurosci.* 8:376. doi: 10.3389/fnins.2014.00376
- Contreras-Vidal, J. L., and Grossman, R. G. (2013). "Neurorex: a clinical neural interface roadmap for EEG-based brain machine interfaces to a lower body robotic exoskeleton," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka: IEEE), 1579–1582.
- Contreras-Vidal, J. L., Kilicarslan, A., Huang, H. H., and Grossman, R. G. (2015). Human-centered design of wearable neuroprostheses and exoskeletons. *AI Magazine* 36, 12–22.
- Cramer, S. C., Lastra, L., Lacourse, M. G., and Cohen, M. J. (2005). Brain motor system function after chronic, complete spinal cord injury. *Brain* 128, 2941–2950. doi: 10.1093/brain/awh648
- Fitzsimmons, N., Lebedev, M., Peikon, I., and Nicolelis, M. A. (2009). Extracting kinematic parameters for monkey bipedal walking from cortical neuronal ensemble activity. *Front. Integr. Neurosci.* 3:3. doi: 10.3389/neuro.07.003.2009
- French, J. A. (2014). *Towards the Implementation of Non-Invasive Brain Machine Interface Control on a Rehabilitative Robotic Upper Limb Exoskeleton*. Master's Thesis, Rice University. Available online at: <http://hdl.handle.net/1911/76725>
- Garrett, D., Peterson, D. A., Anderson, C. W., and Thaut, M. H. (2003). Comparison of linear, nonlinear, and feature selection methods for eeg signal classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 11, 141–144. doi: 10.1109/TNSRE.2003.814441
- Gentili, R. J., Bradberry, T. J., Oh, H., Costanzo, M. E., Kerick, S. E., Contreras-Vidal, J. L., et al. (2015). Evolution of cerebral cortico-cortical communication during visuomotor adaptation to a cognitive-motor executive challenge. *Biol. Psychol.* 105, 51–65. doi: 10.1016/j.biopsycho.2014.12.003
- Giacometti, P., Perdue, K. L., and Diamond, S. G. (2014). Algorithm to find high density EEG scalp coordinates and analysis of their correspondence to structural and functional regions of the brain. *J. Neurosci. Methods* 229, 84–96. doi: 10.1016/j.jneumeth.2014.04.020
- Gobel, E. W., Parrish, T. B., and Reber, P. J. (2011). Neural correlates of skill acquisition: decreased cortical activity during a serial interception sequence learning task. *Neuroimage* 58, 1150–1157. doi: 10.1016/j.neuroimage.2011.06.090
- Gourab, K., and Schmit, B. D. (2010). Changes in movement-related  $\beta$ -band eeg signals in human spinal cord injury. *Clin. Neurophysiol.* 121, 2017–2023. doi: 10.1016/j.clinph.2010.05.012
- Hortal, E., Planelles, D., Costa, A., Iáñez, E., Úbeda, A., Azorín, J., et al. (2015). Svm-based brain-machine interface for controlling a robot arm through four mental tasks. *Neurocomputing* 151, 116–121. doi: 10.1016/j.neucom.2014.09.078
- Jiang, N., Gizzi, L., Mrachacz-Kersting, N., Dremstrup, K., and Farina, D. (2015). A brain-computer interface for single-trial detection of gait initiation from movement related cortical potentials. *Clin. Neurophysiol.* 126, 154–159. doi: 10.1016/j.clinph.2014.05.003
- Jorquera, F. S. M., Grassi, S., Farine, P.-A., and Contreras-Vidal, J. L. (2013). "Classification of stance and swing gait states during treadmill walking from non-invasive scalp electroencephalographic (EEG) signals," in *Converging Clinical and Engineering Research on Neurorehabilitation, Vol. 1, Biosystems & Biorobotics*, eds J. Pons, D. Torricelli, and

- M. Pajaro (Berlin: Springer), 507–511. doi: 10.1007/978-3-642-34546-3\_81
- Kilicarslan, A., Prasad, S., Grossman, R. G., and Contreras-Vidal, J. L. (2013). “High accuracy decoding of user intentions using EEG to control a lower-body exoskeleton,” in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka: IEEE), 5606–5609. doi: 10.1109/embc.2013.6610821
- Kranczoch, C., Athanassiou, S., Shen, S., Gao, G., and Sterr, A. (2008). Short-term learning of a visually guided power-grip task is associated with dynamic changes in eeg oscillatory activity. *Clin. Neurophysiol.* 119, 1419–1430. doi: 10.1016/j.clinph.2008.02.011
- Lal, T. N., Schröder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., et al. (2004). Support vector channel selection in BCI. *IEEE Trans. Biomed. Eng.* 51, 1003–1010. doi: 10.1109/TBME.2004.827827
- Leamy, D. J., Kocijan, J., Domijan, K., Duffin, J., Roche, R. A., Commins, S., et al. (2014). An exploration of EEG features during recovery following stroke—implications for BCI-mediated neurorehabilitation therapy. *J. Neuroeng. Rehabil.* 11:9. doi: 10.1186/1743-0003-11-9
- Leeb, R., Lancelle, M., Kaiser, V., Fellner, D. W., and Pfurtscheller, G. (2013). Thinking penguin: multimodal brain–computer interface control of a vr game. *IEEE Trans. Comput. Intell. AI Games* 5, 117–128. doi: 10.1109/TCIAIG.2013.2242072
- Lew, E. Y. L., Chavarriaga, R., Silvoni, S., and Millán, J. D. R. (2014). Single trial prediction of self-paced reaching directions from EEG signals. *Front. Neurosci.* 8:222. doi: 10.3389/fnins.2014.00222
- Li, X., Chen, X., Yan, Y., Wei, W., and Wang, Z. J. (2014). Classification of eeg signals using a multiple kernel learning support vector machine. *Sensors* 14, 12784–12802. doi: 10.3390/s140712784
- Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *J. Neural Eng.* 4:R1. doi: 10.1088/1741-2560/4/2/R01
- Luu, T. P., He, Y., Brown, S., Nakagame, S., and Contreras-Vidal, J. L. (2016). Gait adaptation to visual kinematic perturbations using a real-time closed-loop brain–computer interface to a virtual reality avatar. *J. Neural Eng.* 13:036006. doi: 10.1088/1741-2560/13/3/036006
- Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., et al. (2013). “Real-time modeling and 3d visualization of source dynamics and connectivity using wearable EEG,” in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (Osaka: IEEE), 2184–2187. doi: 10.1109/embc.2013.6609968
- Nathan, K., and Contreras-Vidal, J. L. (2015). Negligible motion artifacts in scalp electroencephalography (EEG) during treadmill walking. *Front. Hum. Neurosci.* 9:708. doi: 10.3389/fnhum.2015.00708
- Niazi, I. K., Jiang, N., Jochumsen, M., Nielsen, J. F., Dremstrup, K., and Farina, D. (Jan. 2013). Detection of movement-related cortical potentials based on subject-independent training. *Med. Biol. Eng. Comput.* 51, 507–512. doi: 10.1007/s11517-012-1018-1
- Noda, T., Sugimoto, N., Furukawa, J.-I., Sato, M., Hyon, S.-H., and Morimoto, J. (2012). “Brain-controlled exoskeleton robot for bmi rehabilitation,” in *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (Osaka: IEEE), 21–27. doi: 10.1109/HUMANOIDS.2012.6651494
- Presacco, A., Forrester, L. W., and Contreras-Vidal, J. L. (2012). Decoding intra-limb and inter-limb kinematics during treadmill walking from scalp electroencephalographic (EEG) signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* 20, 212–219. doi: 10.1109/TNSRE.2012.2188304
- Presacco, A., Goodman, R., Forrester, L., and Contreras-Vidal, J. L. (2011). Neural decoding of treadmill walking from noninvasive electroencephalographic signals. *J. Neurophysiol.* 106, 1875–1887. doi: 10.1152/jn.00104.2011
- Rakotomamonjy, A. (2008). *SimpleMKL Toolbox*. Available online at: <http://asi.insa-rouen.fr/enseignants/~arakoto/code/mklindex.html>
- Rakotomamonjy, A., Bach, F., Canu, S., and Grandvalet, Y. (2008). SimpleMKL. *J. Mach. Learn. Res.* 9, 2491–2521.
- Sale, P., Franceschini, M., Waldner, A., and Hesse, S. (2012). Use of the robot assisted gait therapy in rehabilitation of patients with stroke and spinal cord injury. *Eur. J. Phys. Rehabil. Med.* 48, 111–121.
- Samek, W., Binder, A., and Muller, K.-R. (2013). “Multiple kernel learning for brain-computer interfacing,” in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (Osaka: IEEE), 7048–7051.
- Sarnthein, J., Stern, J., Aufenberg, C., Rousson, V., and Jeanmonod, D. (2006). Increased EEG power and slowed dominant frequency in patients with neurogenic pain. *Brain* 129, 55–64. doi: 10.1093/brain/awh631
- Shawe-Taylor, N., and Kandola, A. (2002). On kernel target alignment. *Adv. Neural Inf. Proc. Syst.* 14, 367–373.
- Sonnenburg, S., Rätsch, G., Schäfer, C., and Schölkopf, B. (2006). Large scale multiple kernel learning. *J. Mach. Learn. Res.* 7, 1531–1565.
- Tian, X., Gasso, G., and Canu, S. (2012). A multiple kernel framework for inductive semi-supervised SVM learning. *Neurocomputing* 90, 46–58. doi: 10.1016/j.neucom.2011.12.036
- Tran, Y., Boord, P., Middleton, J., and Craig, A. (2004). Levels of brain wave activity (8–13 hz) in persons with spinal cord injury. *Spinal Cord* 42, 73–79. doi: 10.1038/sj.sc.3101543
- Vapnik, V. N., and Vapnik, V. (1998). *Statistical Learning Theory*, vol. 1. New York, NY: Wiley.
- Velu, P., and de Sa, V. R. (2013). Single-trial classification of gait and point movement preparation from human EEG. *Front. Neurosci.* 7:84. doi: 10.3389/fnins.2013.00084
- Venkatakrishnan, A., Francisco, G. E., and Contreras-Vidal, J. L. (2014). Applications of brain–machine interface systems in stroke recovery and rehabilitation. *Curr. Phys. Med. Rehabil. Rep.* 2, 93–105. doi: 10.1007/s40141-014-0051-4
- Wu, J., Stoica, B. A., Luo, T., Sabirzhanov, B., Zhao, Z., Guanciale, K., et al. (2014a). Isolated spinal cord contusion in rats induces chronic brain neuroinflammation, neurodegeneration, and cognitive impairment: involvement of cell cycle activation. *Cell Cycle* 13, 2446–2458. doi: 10.4161/cc.29420
- Wu, J., Zhao, Z., Sabirzhanov, B., Stoica, B. A., Kumar, A., Luo, T., et al. (2014b). Spinal cord injury causes brain inflammation associated with cognitive and affective changes: role of cell cycle pathways. *J. Neurosci.* 34, 10989–11006. doi: 10.1523/JNEUROSCI.5110-13.2014
- Zhang, Y., Yang, H. L., Prasad, S., Pasolli, E., Jung, J., and Crawford, M. (2015). Ensemble multiple kernel active learning for classification of multisource remote sensing data. *IEEE J. Selected Topics Appl. Earth Observ. Remote Sens.* 8, 845–858. doi: 10.1109/JSTARS.2014.2359136

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Zhang, Prasad, Kilicarslan and Contreras-Vidal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.