IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 November 2023, accepted 4 December 2023, date of publication 7 December 2023,
date of current version 14 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3340685

RESEARCH ARTICLE

Common Spatial Pattern and Riemannian Manifold-Based Real-Time Multiclass Motor Imagery EEG Classification

KUO-KAI SHYU, (Senior Member, IEEE), SZU-CHI HUANG, KAI-JEN TUNG, LUNG-HAO LEE, (Member, IEEE), PO-LEI LEE, (Member, IEEE), AND YU-HAO CHEN
Department of Electrical Engineering, National Central University, Taoyuan 32001, Taiwan

Corresponding author: Szu-Chi Huang (r310102930719@gmail.com)

This work was supported in part by the National Science and Technology Council, Taiwan, under Grant MOST 111-2221-E-008-101-MY3.

ABSTRACT Several motor imagery classification methods have been developed and achieve higher accuracy. Machine learning (ML) based algorithms utilizing manually designed features often encounter robustness issues, leading to diminished accuracy. While deep learning (DL) based algorithms exhibit promising accuracy, their extensive computational requirements present challenges in implementing them on portable devices, thereby restricting their practical applications. In this paper, we improve the ML-based algorithm's feature robustness problems by combining common spatial patterns with Riemannian tangent space mapping, enhancing the algorithm's feature quality. Furthermore, we introduce a method that utilizes the distance between data points and the SVM hyperplane to compute category scores, thereby enhancing classifier performance. Our experiment uses the BCI Competition IV 2A, BCI Competition III 3A, and a self-recorded dataset for subject-specific experiments to validate the algorithm's classification performance. Experimental results show that the proposed algorithm achieves the best classification performance, with an accuracy of 78.55%, 83.33%, and 57.44% for BCI Competition IV 2A, BCI Competition III 3A, and the self-recorded dataset. Additionally, to assess the practicality of a real-time portable application, we implemented the proposed algorithm on Raspberry Pi and Jetson Nano, measuring their computation time and peak memory usage. The results demonstrate that our algorithm necessitates only 0.08 to 0.3 seconds of computation time and employs a mere 15MB of memory.

INDEX TERMS Electroencephalograph, motor imagery, common spatial pattern, Riemannian tangent space, filter banks, support vector machine.

I. INTRODUCTION

A Brain-Computer Interface (BCI) bridges the brain and external devices by translating brain signals into corresponding commands. BCI system can assist the rehabilitation process of paralyzed patients and improve their self-care ability. Taking stroke patients as an example, stroke highly likely results in varying degrees of limb impairment in patients [1]. Commonly employed clinical rehabilitation methods include physiotherapy (PT), occupational therapy (OT), robot therapy, electrical stimulation, pharmacological therapy, and virtual reality-assisted motor imagery therapy [2]. However, for patients with severe limb paralysis, physical rehabilitation treatments (such as PT and OT) are difficult to perform [3]. In contrast, motor imagery does not require physical movement and can facilitate brain neural network reorganization, thereby promoting impaired brain function recovery [2], [4].

Currently, intelligent BCI systems provide real-time feedback based on physiological signals such as Positron Emission Tomography (PET), Functional Magnetic Resonance Imaging (fMRI), Near-Infrared Spectroscopy (fNIRS), Magnetoencephalography (MEG), and Electroencephalography (EEG). PET, fNIRS, and fMRI indirectly measure brain activity with lower temporal resolution. Systems utilizing

The associate editor coordinating the review of this manuscript and approving it for publication was Sangsoon Lim.

Electromyography (EMG) are bulkier and excessively costly, making them impractical for real-world applications [2]. However, EEG-based systems offer high temporal resolution and cost-effectiveness and are widely employed in BCI systems [5], [6], [7].

In EEG-based BCI systems, EEG motor imagery classification algorithms can link the EEG signal and device command. ML-based motor imagery classification algorithms' architecture can be divided into three parts: preprocessing, feature extraction, and classification. Among these, feature extraction plays a pivotal role in the algorithm. Commonly utilized signal features encompass Fourier Transform, Event-Related Synchronization (ERS), Power Spectral Density (PSD), and Common Spatial Patterns (CSP). Nevertheless, features based on the Fourier Transform tend to ignore EEG's temporal pattern [7]. Features based on ERS and PSD are susceptible to electrode channels [6], [7]. CSP-based features are highly dependent on operation frequency bands [8]. Consequently, manually engineered features also suffer from robustness issues [6], [9].

In the pursuit of achieving more accurate classification performance, DL-based algorithms have been explored. For instance, Zhang et al. [6] employed graph embedding to represent spatial nodes of EEG signals, followed by CNN for extracting spatio-temporal features. Subsequently, classification was conducted through attention layers and dense layers. Yang et al. [10] utilized Riemannian covariance as a feature and multi-layer perceptron for classification. However, algorithms based on deep learning often demand higher computational resources, rendering them challenging to implement on portable devices [2], [11], thus imposing limitations on practical applications.

This paper proposed the Discriminative Filter Bank Tangent Space Mapping and Common Spatial Pattern (DFBTSM-CSP) algorithm, which integrates Riemannian tangent space and CSP for EEG feature extraction, employing SVM as the classifier. During classification, we calculate category scores based on the distance between data points and the SVM hyperplane, followed by classification using these category scores. The main contributions of this paper and key advantages of the DFBTSM-CSP algorithm are summarized as follows:

- We integrate CSP and Riemannian tangent space for feature extraction. CSP algorithm using optimal spatial filters maximizes inter-class variance, while Riemannian-based feature projection of EEG spatial covariance matrices onto the Riemannian tangent space enhances discriminability and robustness. Experimental results show that the proposed method achieves more accuracy due to the feature providing a more comprehensive analysis of EEG spatial information.
- The proposed methods' final classification decisions are based on category scores computed using the distance between data points and the SVM hyperplane. Experimental results demonstrate that utilizing category scores for classification decisions outperforms traditional SVM, yielding superior classification performance.
- We implement frequency band selection using the Fisher ratio, reducing the number of frequency bands by approximately 20-40% and concurrently improving algorithm accuracy by around 2%.
- We validate the feasibility of real-time wearable applications by implementing the algorithm on an embedded system. The results reveal that the proposed algorithm requires a peak memory usage of only 14.85MB and a computation time of merely 0.2 seconds.

The remaining sections of this paper are as follows. Section II provides details of the proposed method. Section III introduces the dataset and the recording method. We present the experimental results and discussions in Section IV. In Section V, we summarize the contributions of this paper and outline algorithm limitations.

II. METHODOLOGY

The proposed DFBTSM-CSP classification algorithm is shown in Fig. 1. The algorithm architecture can be divided into four parts, including preprocessing, filter bank and band selection, feature extraction, and classification. In the preprocessing phase, noise in the EEG signal is eliminated while retaining frequency bands relevant to motor imagery through the use of low-pass and band-pass filters. During the frequency band selection phase, the EEG signal is decomposed into multiple sub-bands, and the most discriminative frequency bands are chosen based on the training data for feature extraction. This paper employs CSP and Riemannian tangent space as signal features, followed by the computation of category scores based on the distance between the features and the SVM hyperplane to determine the classification outcomes.

A. DATA PREPROCESSING

The frequency range of motor imagery EEG signals is between 1 and 50Hz. Since the 60Hz power-line noise amplitude is much larger than the EEG signal, the power-line noise cannot be effectively removed with a band-pass filter. Therefore, a band-stop filter of 55 to 65 Hz is used to remove the 60Hz noise, and a band-pass filter of 1 to 50 Hz is used to filter out the unrelative signal.

B. FILTER BANK AND BAND SELECTION

The most discriminative frequency band of each subject is not the same [12], [13], so it is crucial to select the suitable frequency band. In this part, the EEG signal is split into different frequency bands through multiple band-pass filters, and the Fisher ratio band selection algorithm is used to select the most discriminative band. In this paper, frequency band selection was conducted using multiple narrowband and broadband sub-bands within the 4 to 35 Hz frequency range. Narrowband signals encompass more prominent motor imagery

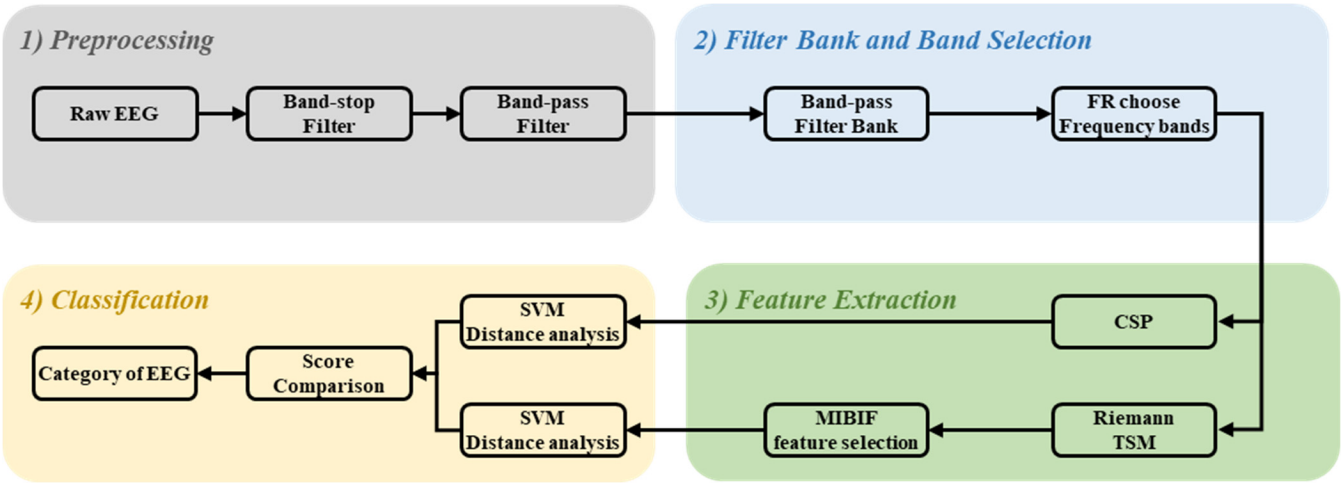

FIGURE 1. The proposed DFBTSM-CSP classification algorithm.

information and are less susceptible to noise interference, while broadband signals include some apparent features that can enhance motor imagery performance [14]. Employing frequency band selection also facilitates the removal of less discriminative frequency bands. The band-pass filter bands of the filter banks used in this paper are respectively 4-10Hz, 7-13Hz, 11-17Hz, 15-21Hz, 18-24Hz, 21-27Hz, 24-31Hz, 27-34Hz, 31-37Hz, 4-14Hz, 13-23Hz, 22-32Hz, and 8-35Hz.

The Fisher ratio band selection algorithm [12], [15] ranks the bands according to inter-class variance and intra-class variance of the band power, and selects the most discriminative band. Firstly, the power of specific frequency band in each trial is calculated, which can be expressed as (1).

$$P_{t,f} = \frac{1}{N}\sum_{n=1}^{N} x_{t,f}(n)^2 \tag{1}$$

where $P_{t,f}$ is the power of the *f-th* frequency band of the *t-th* test, N is the total number of sampling points, $x_{t,f}(n)$ is the signal of the *t-th* sampling point of the *f-th* frequency band.

Then one calculates the average power of each frequency band and the average power of the specific frequency band of each category, as shown in (2) and (3) respectively.

$$m_f = \frac{\sum_{t=1}^{M} P_{t,f}}{T} \tag{2}$$

$$m_{f,c} = \frac{\sum_{t_c=1}^{M_c} P_{t_c,f}}{T_c} \tag{3}$$

It is noticed that m_f in (2) is the average power of the *f-th* frequency band, t is the number of trials, M is the total number of trials. In (3), $m_{f,c}$ is the average power of the *f-th* frequency band of class C, t_c is the trial of class C, M_c is the total number of trials of class C, $P_{t_c,f}$ is the power of the *f-th* frequency band of the trials of class C.

Accordingly, the intra-class variation and inter-class variation values under specific frequency band are calculated, as shown in (4) and (5) respectively.

$$S_{W,f} = \sum_{c=1}^{D}\sum_{t_c=1}^{M_c}\left(P_{t_c,f} - m_{f,c}\right)^2 \tag{4}$$

$$S_{B,f} = \sum_{c=1}^{D} M_c\left(m_f - m_{f,c}\right)^2 \tag{5}$$

Note that $S_{W,f}$ in (4) is the intra-class variation value of the *f-th* frequency band, D is the total number of classes, C is the class. And $S_{B,f}$ is the inter-class variation value of the *f-th* frequency band. Therefore, the Fisher ratio $F_{R,f}$ can be obtained to evaluate the importance of the frequency band. The calculation of $F_{R,f}$ is given as.

$$F_{R,f} = \frac{S_{B,f}}{S_{W,f}} \tag{6}$$

where $F_{R,f}$ is the FR value of the *f-th* frequency band. The higher the FR value, the more discriminative the frequency band is. Therefore, the importance of each frequency band can be ranked according to the FR value.

C. FEATURE EXTRACTION

1) TANGENT SPACE MAPPING OF RIEMANNIAN MANIFOLDS

Compared with Euclidean space, Riemannian space can more accurately describe the correlation between high-dimensional EEG signals. This property may achieve good results in the classification of motor imagery signals [16], [17].

Firstly, the sample covariance matrix is calculated as

$$P_t = \frac{1}{N-1} X_t X_t^T \tag{7}$$

where $P_t \in \mathbb{R}^{N\times N}$ is the sample covariance matrix of the *t-th* test, N is the total number of sampling points, superscript T is the sign of transpose matrix, X_t is the EEG signal of the t_{th} test.

Then calculate the Euclidean mean by Eq. (8) as the initial value of the Riemann mean.

$$P_E = \frac{1}{M}\sum_{t=1}^{M} P_t \tag{8}$$

where P_E is the Euclidean mean value and M is the total number of tests. Note that because the calculation of the Riemannian mean P_R is not an analytical solution, it needs to be approximated by iterative method.

We can project the point P_i from Reimann to tangent space by performing the Riemann Log map operation through Eq. (9). And perform the Riemann Exp map operation through Eq. (10) to project the point S_i from the tangent space to the Riemann space.

$$Log_P(P_i) = P^{\frac{1}{2}} Log(P^{-\frac{1}{2}} P_i P^{-\frac{1}{2}}) P^{-\frac{1}{2}} \tag{9}$$

$$Exp_P(S_i) = P^{\frac{1}{2}} Exp(P^{-\frac{1}{2}} S_i P^{-\frac{1}{2}}) P^{-\frac{1}{2}} \tag{10}$$

The Riemann mean algorithm [16] performs the Riemann Log map operation on all training data and calculates the average value of the tangent space S, and then performs the Riemann Exp map operation on S to obtain the Riemann mean value P_R, and iterates until the change in P_R is less than the tolerance ε. The Riemann mean algorithm is summarized in Fig. 2.

Algorithm 1: Riemann mean algorithm

Input: covariance matrix of EEG $P_1, P_2, P_3, \ldots, P_N$
tolerance ε

Output: Reimann mean P_R

/* Set initial Reimann mean P_R as Euclidean mean P_E */

$P_R^0 \leftarrow \frac{1}{M}\sum_{t=1}^{M} P_t$

$i \leftarrow 0$

Do

  /* calculate mean on the tangent space */

  $S \leftarrow \frac{1}{M}\sum_{t=1}^{M} Log_{P_R^i}(P_t)$

  /* mapping S to Riemannian manifold */

  $P_R^{i+1} \leftarrow \mathrm{Exp}_{\mathrm{P_R^i}}(S)$

  $i \leftarrow i + 1$

While $\left\| P_R^i - P_R^{i-1} \right\|_F > \varepsilon$

FIGURE 2. Pseudocode of Riemann mean algorithm.

After obtaining the Riemannian mean, the Riemannian manifold can map to the tangent space by Eq. (9). Thus, the Euclidean space vector [18] is extracted from the tangent space $T_{s,t}$ as a feature, and the extraction method can be expressed as

$$v_t = upper\left(T_{s,t}\right) = upper\left(Log_{P_R}\left(P_t\right)\right) \tag{11}$$

where v_t is the Euclidean space vector extracted from tangent space $T_{s,t}$, $upper(\cdot)$ means the upper triangular part of the matrix, and the off-diagonal elements are multiplied by the weight $\sqrt{2}$. Note that the upper triangular elements are converted into a one-dimensional vector.

The extracted v_t vector is used to select the better features using Mutual Information based Best Individual Feature (MIBIF) [19], [20].

Suppose there is a set of labels Y, and a set of feature vectors F, each of which contains d features, where $d = \frac{C_h(C_h+1)}{2}$, $F = \{f_1, f_2, \ldots f_d\}$, we need to find out the better k features from the d features by the score of mutual information. The mutual information formula is

$$I\left(f_i; Y\right) = H\left(f_i\right) - H\left(f_i \mid Y\right) \tag{12}$$

where $I\left(f_i; Y\right)$ is the mutual information between f_i and Y, $i = 1, 2, \ldots d$, $H\left(f_i\right)$ is the entropy of f_i, and $H\left(f_i \mid Y\right)$ is the conditional entropy of f_i given Y. $H\left(f_i\right)$ and $H\left(f_i \mid Y\right)$ are respectively given as (13) and (14).

$$H\left(f_i\right) = -\sum_{f_i \in F} p\left(f_i\right) log_2 p\left(f_i\right) \tag{13}$$

$$H\left(f_i \mid Y\right) = -\sum_{y \in Y}\sum_{f_i \in F} p(y, f_i) log_2 p(f_i|y) \tag{14}$$

where $p\left(\cdot\right)$ is the probability function. Through (12) to (14), one gets d number of the mutual message scores. Then k features with higher scores are chosen as the selected features

2) COMMON SPATIAL PATTERNS

Common Spatial Patterns [21] is a way to maximize the difference in the variation of two categories time-domain signals. From multiple electrode channels, it can find a set of optimal spatial filter. The inner product of the spatial filter and time domain signals can make the variation of numerical difference between the different categories to maximize, and the variation of numerical differences between the same category to minimize.

Firstly, the normalized covariance matrices of time domain signals of the two categories are calculated as.

$$E_1 = \frac{X_1 \cdot X_1^T}{trace(X_1 \cdot X_1^T)}, E_2 = \frac{X_2 \cdot X_2^T}{trace\left(X_2 \cdot X_2^T\right)} \tag{15}$$

where E_1 and E_2 are respectively the normalized covariance matrices of category 1 and category 2; X_1 and X_2 are respectively the time domain signals of category 1 and category 2; and *trace* is the trace of the calculation matrix. Then, the covariance matrix of the mixed space is given as

$$E = \overline{E_1} + \overline{E_2} \tag{16}$$

where E is the covariance matrix of the mixed space. $\overline{E_1}$ and $\overline{E_2}$ are the average normalized covariance matrices of category 1 and category 2, respectively. The eigenvalue decomposition of the covariance matrix in the mixed space is done with the eigen-diagonal matrix arranged in descending order, as shown in (17).

$$E = U\lambda U^T \tag{17}$$

where U is the eigenvector matrix of matrix E and λ is the eigen diagonal matrix. Then, the whitening matrix can be obtained as.

$$P = \sqrt{\lambda}^{-1} \cdot U^T \tag{18}$$

where P is the whitening matrix. Then perform a whitening transformation between $\overline{E_1}$ and $\overline{E_2}$ as (19).

$$S_1 = P \cdot \overline{E_1} \cdot P^T, \quad S_2 = P \cdot \overline{E_2} \cdot P^T \tag{19}$$

where S_1 and S_2 are respectively the average normalized covariance matrix of category 1 and category 2. The eigenvalue decompositions of S_1 and S_2 respectively have the eigen diagonal matrix of S_1 arranged in descending order, and that matrix of S_2 arranged in ascending order, as shown in (20).

$$S_1 = B_1 \cdot \lambda_1 \cdot B_1^T, \quad S_2 = B_2 \cdot \lambda_2 \cdot B_2^T \tag{20}$$

It is noticed that $B_1 = B_2$, so the eigenvector matrix of S_1 is the same as S_2, and the sum of eigen diagonal matrices λ_1 and λ_2 is the identity matrix as

$$\lambda_1 + \lambda_2 = I \tag{21}$$

Since the sum of the eigenvalues of the two categories is always one, the eigenvector corresponding to the maximum eigenvalue of S_1 will also correspond to the minimum eigenvalue of S_2, and vice versa. Therefore, a spatial filter can be obtained to maximize the difference of variation values of different categories and minimize the difference of variation values of the same category.

The spatial filter is given as

$$W = B_1^T P \tag{22}$$

The spatial projection matrix W is the desired spatial filter. Then one has the inner product of the spatial filter W with the original brain signal X as

$$Z = W \cdot X \tag{23}$$

where Z is the spatially filtered signal. Thus, the first and last m column vectors of Z are used to calculate the variation values of the column vectors.

Accordingly, $2m$ features can be obtained as

$$f_p = \log\left(\frac{var\left(Z_p\right)}{\sum_{q=1}^{2m} var\left(Z_q\right)}\right), p = 1, 2 \ldots, 2m \tag{24}$$

where f_p is the feature obtained by taking out the p_{-th} column vector, and $var(\cdot)$ is the variation value.

D. CLASSIFICATION

In the feature extraction, we use a one-versus-one support vector machine [22] to train multiple binary classifiers to vote for the multi-classification task. We respectively perform SVM on the features from Reimann TSM and CSP. Then calculate the category scores through Eq. (25) and (26). Finally, the result with a higher category score is used as the classification result. The classification process of SVM distance analysis is shown in Fig. 3.

$$L_x = \frac{1}{\|w\|} |w \cdot f + b| \tag{25}$$

$$score = \sum_{SVM_x(f) \in C} \frac{L_x}{M_x} \tag{26}$$

where w is the normal vector of the hyperplane, f is the input feature vector, b is the offset of the hyperplane relative to the origin, L_x is the distance between the test point and the hyperplane, M_x is the average distance between the training points and the hyperplane, and C is the predict category.

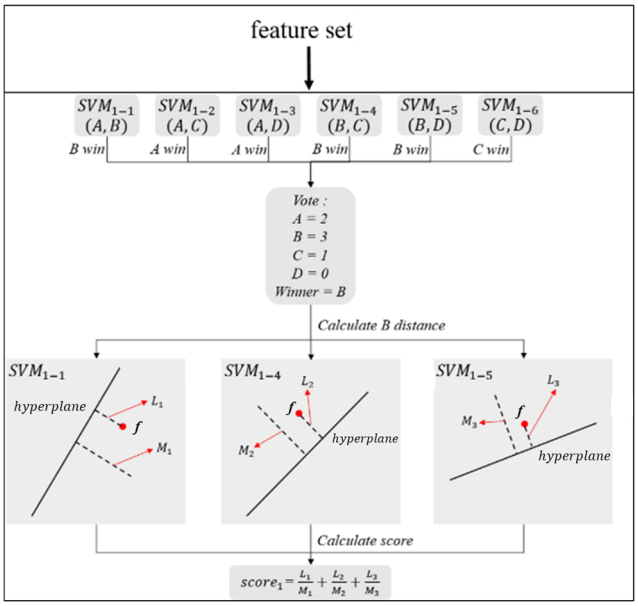

FIGURE 3. The process flow of SVM distance analysis classification.

III. DATASET

This paper conducts experiments using two publicly available datasets (BCI Competition IV 2A and III 3A) along with a self-recorded dataset to validate the algorithm's performance through subject-specific experiments. The training and testing data of the datasets employed in this study were recorded from subjects on different days. In BCI Competition IV 2A dataset [23], there are 9 subjects containing four categories of motor imagery (left hand, right hand, feet and tongue). Each category contains 72 training sets and 72 testing sets, so there are 288 trials in the training set and the testing set The EEG signal was sampled at 250Hz and had 22 electrode channels.

BCI Competition III 3A dataset [24] has 3 subjects and contains four categories of motor imaginary, (left hand, right hand, feet and tongue). The first subject (K3B) contains 90 trials in each category and a total of 360 trials, among which 180 trials are used as training set and the remaining 180 trials as testing set. As for the second subject (K6B) and third subject (L1B), each category contained 60 trials, a total of 240 trials, of which 120 trials were used as the training set and the remaining 120 trials as the test set. The brain signal was sampled at 250Hz with 60 electrode channels. Due to the excessive number of electrode channels, which greatly increases the computational burden, this study only uses electrode channels in the motor and somatosensory areas of the cerebral cortex, including electrode channels from 17 to 45, as shown in Fig. 4.

The self-recorded motor imaginary brain signal dataset includes five right-handed men, aged 23 or 24, who sat in a moderately high armchair, rested their hands on their legs when not moving, and did not blink during imaginary exercise. The data set consists of four categories, namely left hand, right hand, both hands or right foot and no action, which

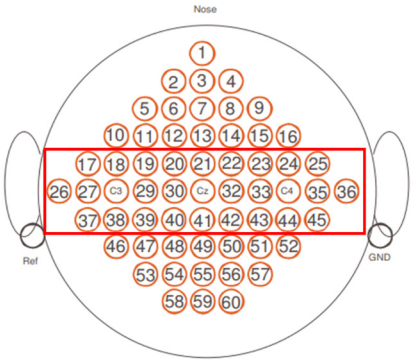

FIGURE 4. BCI competition III 3a EEG electrodes.

are recorded in four periods on the same day. The first two periods are used as the training set and the last two periods are used as the test set. There are 40-80 trials in the training set and 40-80 trials in the test set. The sampling frequency of brain signal is 1000Hz, and there are 8 electrode channels. The electrode channel positions are shown in Fig. 5.

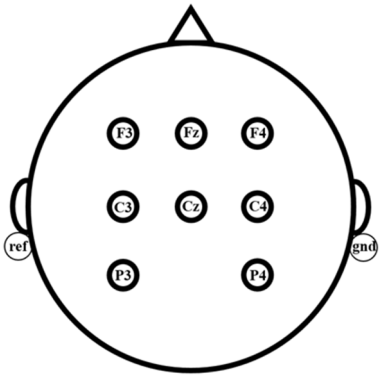

FIGURE 5. Self-recorded EEG electrodes.

The trial recording process is shown in Fig. 6. At the beginning of the test (t=0s), the screen goes blank; at 3 seconds, there are text and picture occur (t=3s); at 5 seconds, there are hints for motor imagery (t=5∼9s); at 9 seconds, the screen goes blank (t=9s). Then rest and wait for the next trial to begin.

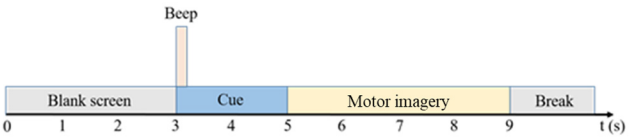

FIGURE 6. Self-recorded EEG recording process.

IV. RESULT

This paper implements the algorithm using Python, and the experimental workflow is illustrated in Fig. 7. In assessing the algorithm's resource usage, we train the proposed model on a personal computer (PC) and subsequently deploy the model on embedded systems (Jetson Nano or Raspberry Pi) to evaluate computational time and memory usage. A subject-specific approach is employed to evaluate the algorithm's classification performance. In order to prevent the data leakage problem, training and testing data are recorded on different days, which is similar to the experimental procedures of prior studies [5], [25], [26], [27].

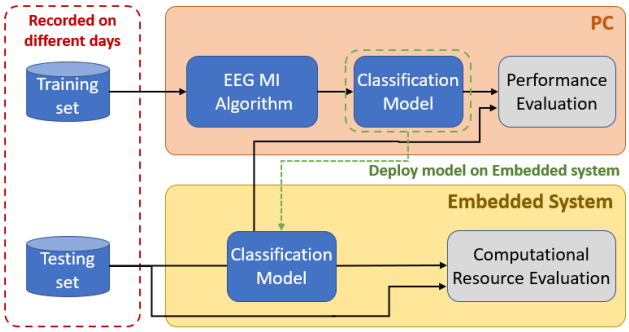

FIGURE 7. Experimental flow chart.

A. PERFORMANCE COMPARISON WITH PREVIOUS ALGORITHM

We compared the proposed method with CSP [21], FBCSP [19], TSM [17], FBTSM [18], FBTSM-MLP [10], FBCSP-PLV [28], CTFSP [29], and BECSP [30]. To further validate the efficacy of our proposed approaches, we included Proposed Method 1 and Proposed Method 2 in the comparison. Proposed Method 1 omits the FR frequency selection in DFBTSM-CSP and directly employs SVM for classification. Proposed Method 2 removes the FR frequency selection in DFBTSM-CSP.

As shown in Table 1, the results indicate that the method proposed exhibits superior performance, with accuracies of 78.55%, 83.33%, and 57.44% on the BCI Competition IV 2A dataset, BCI Competition III 3A dataset, and the self-recorded dataset, respectively. Proposed Method 1 outperforms other methods, indicating that the proposed feature is more discriminative. Comparison between Proposed Method 1 and Proposed Method 2 shows that employing SVM distance-based category score computation yields better classification performance. Furthermore, contrasting Proposed Method 2 with DFBTSM-CSP reveals that removing non-discriminative data through frequency band selection reduces feature dimension and training noise, resulting in improved classification performance.

B. EFFECT ANALYSIS OF FREQUENCY BAND SELECTION

Table 2 and Table 3 make a comparison of proposed method 2 (without band selection) and DFBTSM-CSP (with band selection), and show the effectiveness of the band selection. Table 2 shows the number of frequency bands used is reduced by 19.16% on the BCI competition IV 2a dataset, 42.61% on the BCI competition IV 3a dataset, and 20% on

TABLE 1. Comparison of motor imagery classification algorithm performance.

Published Work	Classification Method	BCI IV 2a dataset mean accuracy	BCI IV 3a dataset mean accuracy	Self-recorded dataset mean accuracy
Ramoser et.al (2000) [21]	CSP + SVM	0.6443	0.7472	0.4587
Keng et.al. (2008) [19]	FBCSP + SVM	0.7041	0.7833	0.4943
Barachant et.al. (2012) [17]	TSM + SVM	0.6875	0.7527	0.4371
Islam et.al. (2018) [18]	FBTSM + SVM	0.7070	0.7527	0.5146
Yang et.al. (2020) [10]	FBTSM + MLP	0.7237	0.7731	0.4375
Wang et.al. (2020) [28]	FBCSP-PLV + MK-RVM	0.7179	0.7676	0.5121
Miao et.al. (2021) [29]	CTFSP + SVM	0.6292	0.7139	0.3957
Hou et.al. (2022) [30]	BECSP + SVM	0.7011	0.7981	0.4758
Proposed Method	Method 1 (without SVM distance analysis and band selection)	0.7532	0.8176	0.4952
	Method 2 (without band selection)	0.7682	0.8176	0.5586
	DFBTSM-CSP + SVM	0.7855	0.8333	0.5744

TABLE 2. Number of bands comparison of the proposed method 2 and the DFBTSM-CSP.

	Proposed Method 2	DFBTSM-CSP	Percentage reduction
BCI competition IV 2a dataset	18	14.55	19.16%
BCI competition IV 3a dataset	18	10.33	42.61%
Self-recorded motor imagery dataset	13	10.40	20%

TABLE 3. Classification accuracy comparison of the proposed method 2 and the DFBTSM-CSP.

	Proposed Method 2	DFBTSM -CSP	Improvement
BCI competition IV 2a dataset	76.82%	78.55%	2.25%
BCI competition IV 3a dataset	81.76%	83.33%	1.92%
Self-recorded motor imagery dataset	55.86%	56.94%	1.93%

the self-recorded dataset. Table 3 shows the classification accuracy increase by 2.25% on the BCI competition IV 2a dataset, 1.92% on the BCI competition IV 3a dataset, and 1.93% on the self-recorded dataset.

Summarizing the results of Table 2 and Table 3, the proposed DFBTSM-CSP architecture uses fewer computing resources and achieves better classification performance.

C. EFFECT OF SELECTING BAND IN EMBEDDED SYSTEM

The proposed DFBTSM-CSP architecture also has excellent performance in real-time classification. For brain signal measurement with 8 electrode channels, 1000Hz sampling frequency, sampling point of 3 bytes and the time duration of 4 seconds, the classification time and memory peak are 0.016∼0.026 seconds and 13.83MB respectively on a PC with i5-9500 CPU and DDR4 2666MHz 16GB RAM.

The same experiment on Jetson Nano 2GB has classification time and memory peak as 0.077∼0.122 seconds and 14.854MB respectively. Moreover, the classification time and memory peak respectively are 0.2∼0.307 seconds and 14.851MB peak memory on Raspberry Pi 3B+. These experimental results show that the proposed classification algorithm implemented in real-time applications is feasible.

V. CONCLUSION

This paper integrates CSP and Riemannian tangent space as features and proposes the use of SVM distance analysis for classification. Through subject-specific experiments, this study validates the proposed DFBTSM-CSP algorithm on the BCI Competition IV 2A dataset, BCI Competition III 3A dataset, and a self-recorded dataset, achieving the highest accuracies of 78.55%, 83.33%, and 57.44%, respectively. To further investigate the impact of different components on algorithm performance, we compare Proposed Method 1 (without SVM distance analysis and band selection) and Proposed Method 2 (without band selection). The experimental results of Proposed Method 1 demonstrate that the features introduced in this paper are more discriminative compared to other methods. The results of Proposed Method 2 show that both SVM distance analysis and frequency band selection

enhance the classifier's performance. Further comparison between Proposed Method 2 and DFBTSM-CSP reveals that frequency band selection leads to a 1-2% improvement in accuracy and reduces the use of frequency bands by 20-40%, making the algorithm more feasible for portable devices. This paper also implements the algorithm on embedded systems (Raspberry Pi and Jetson Nano) to verify practicality for portable applications. The results indicate that the computational time of the algorithm is only 0.08 to 0.3 seconds, and it utilizes approximately 15MB of memory.

While this paper focuses on implementing a multi-class motor imagery classification algorithm on portable devices, it does not consider cross-subject application scenarios. In the future, we hope to address the issue of reduced algorithm classification performance due to variations in signal characteristics among different subjects by integrating domain-adaptive algorithms [31].

Furthermore, this paper does not explicitly address crucial artifacts like eye-blinking noise. Currently, the algorithm relies on frequency band selection and uses multiple features to enhance robustness [12], [14], [15], [28], [32]. In future works, there is an aspiration to integrate the algorithm with pertinent techniques [33], [34], [35] This collaborative approach is anticipated to provide a more comprehensive solution to mitigate the effects of intentional disturbances.

REFERENCES

[1] M. Katan and A. Luft, "Global burden of stroke," *Seminars Neurol.*, vol. 38, no. 2, pp. 208–211, Apr. 2018, doi: 10.1055/s-0038-1649503.

[2] C. S. Choy, S. L. Cloherty, E. Pirogova, and Q. Fang, "Virtual reality assisted motor imagery for early post-stroke recovery: A review," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 487–498, 2023, doi: 10.1109/RBME.2022.3165062.

[3] G. Saposnik and M. Levin, "Virtual reality in stroke rehabilitation: A meta-analysis and implications for clinicians," *Stroke*, vol. 42, no. 5, pp. 1380–1386, May 2011, doi: 10.1161/strokeaha.110.605451.

[4] J. Munzert, B. Lorey, and K. Zentgraf, "Cognitive motor processes: The role of motor imagery in the study of motor representations," *Brain Res. Rev.*, vol. 60, no. 2, pp. 306–326, May 2009, doi: 10.1016/j.brainresrev.2008.12.024.

[5] H. Fang, J. Jin, I. Daly, and X. Wang, "Feature extraction method based on filter banks and Riemannian tangent space in motor-imagery BCI," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 6, pp. 2504–2514, Jun. 2022, doi: 10.1109/JBHI.2022.3146274.

[6] D. Zhang, K. Chen, D. Jian, and L. Yao, "Motor imagery classification via temporal attention cues of graph embedded EEG signals," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 9, pp. 2570–2579, Sep. 2020, doi: 10.1109/JBHI.2020.2967128.

[7] M. T. Sadiq, X. Yu, Z. Yuan, M. Z. Aziz, S. Siuly, and W. Ding, "A matrix determinant feature extraction approach for decoding motor and mental imagery EEG in subject-specific tasks," *IEEE Trans. Cognit. Develop. Syst.*, vol. 14, no. 2, pp. 375–387, Jun. 2022, doi: 10.1109/TCDS.2020.3040438.

[8] S.-H. Park, D. Lee, and S.-G. Lee, "Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 498–505, Feb. 2018, doi: 10.1109/TNSRE.2017.2757519.

[9] Y. Li, Y. Liu, W.-G. Cui, Y.-Z. Guo, H. Huang, and Z.-Y. Hu, "Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 4, pp. 782–794, Apr. 2020, doi: 10.1109/TNSRE.2020.2973434.

[10] P. Yang, J. Wang, H. Zhao, and R. Li, "MLP with Riemannian covariance for motor imagery based EEG analysis," *IEEE Access*, vol. 8, pp. 139974–139982, 2020, doi: 10.1109/ACCESS.2020.3011969.

[11] M. Yazid, F. Fahmi, E. Sutanto, W. Shalannanda, R. Shoalihin, G.-J. Horng, and Aripriharta, "Simple detection of epilepsy from EEG signal using local binary pattern transition histogram," *IEEE Access*, vol. 9, pp. 150252–150267, 2021, doi: 10.1109/ACCESS.2021.3126065.

[12] K. P. Thomas, C. Guan, C. T. Lau, A. P. Vinod, and K. K. Ang, "A new discriminative common spatial pattern method for motor imagery brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 11, pp. 2730–2733, Nov. 2009, doi: 10.1109/TBME.2009.2026181.

[13] S. Kumar, A. Sharma, and T. Tsunoda, "An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information," *BMC Bioinf.*, vol. 18, no. S16, p. 545, Dec. 2017, doi: 10.1186/s12859-017-1964-6.

[14] M. K. I. Molla, A. A. Shiam, M. R. Islam, and T. Tanaka, "Discriminative feature selection-based motor imagery classification using EEG signal," *IEEE Access*, vol. 8, pp. 98255–98265, 2020, doi: 10.1109/ACCESS.2020.2996685.

[15] Y.-H. Liu, S. Huang, and Y.-D. Huang, "Motor imagery EEG classification for patients with amyotrophic lateral sclerosis using fractal dimension and Fisher's criterion-based channel selection," *Sensors*, vol. 17, no. 7, p. 1557, Jul. 2017.

[16] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian geometry applied to BCI classification," in *Latent Variable Analysis and Signal Separation*, V. Vigneron, V. Zarzoso, E. Moreau, R. Gribonval, and E. Vincent, Eds. Berlin, Germany: Springer, 2010, pp. 629–636.

[17] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain–computer interface classification by Riemannian geometry," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 4, pp. 920–928, Apr. 2012, doi: 10.1109/TBME.2011.2172210.

[18] M. R. Islam, T. Tanaka, and M. K. I. Molla, "Multiband tangent space mapping and feature selection for classification of EEG during motor imagery," *J. Neural Eng.*, vol. 15, no. 4, Aug. 2018, Art. no. 046021, doi: 10.1088/1741-2552/aac313.

[19] K. Keng Ang, Z. Yang Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 2390–2397, doi: 10.1109/ijcnn.2008.4634130.

[20] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 6, Jun. 2004, Art. no. 066138, doi: 10.1103/physreve.69.066138.

[21] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000, doi: 10.1109/86.895946.

[22] A. Rocha and S. K. Goldenstein, "Multiclass from binary: Expanding one-versus-all, one-versus-one and ECOC-based approaches," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 289–302, Feb. 2014, doi: 10.1109/TNNLS.2013.2274735.

[23] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI competition 2008—Graz data set A," in *Institute for Knowledge Discovery* (Laboratory of Brain-Computer Interfaces), vol. 16. Graz Univ. of Technology, 2008, pp. 1–6.

[24] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. R. Millan, M. Schroder, and N. Birbaumer, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 153–159, Jun. 2006, doi: 10.1109/tnsre.2006.875642.

[25] S. U. Amin, H. Altaheri, G. Muhammad, W. Abdul, and M. Alsulaiman, "Attention-inception and long-short-term memory-based electroencephalography classification for motor imagery tasks in rehabilitation," *IEEE Trans. Ind. Informat.*, vol. 18, no. 8, pp. 5412–5421, Aug. 2022, doi: 10.1109/TII.2021.3132340.

[26] D. M. Hermosilla, R. T. Codorniú, R. L. Baracaldo, R. S. Zamora, D. D. Rodriguez, Y. L. Albuerne, and J. R. N. Álvarez, "Shallow convolutional network excel for classifying motor imagery EEG in BCI applications," *IEEE Access*, vol. 9, pp. 98275–98286, 2021, doi: 10.1109/ACCESS.2021.3091399.

[27] J.-S. Bang, M.-H. Lee, S. Fazli, C. Guan, and S.-W. Lee, "Spatio-spectral feature representation for motor imagery classification using convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 7, pp. 3038–3049, Jul. 2022, doi: 10.1109/TNNLS.2020.3048385.

[28] H. Wang, T. Xu, C. Tang, H. Yue, C. Chen, L. Xu, Z. Pei, J. Dong, A. Bezerianos, and J. Li, "Diverse feature blend based on filter-bank common spatial pattern and brain functional connectivity for multiple motor imagery detection," *IEEE Access*, vol. 8, pp. 155590–155601, 2020, doi: 10.1109/ACCESS.2020.3018962.

[29] Y. Miao, J. Jin, I. Daly, C. Zuo, X. Wang, A. Cichocki, and T.-P. Jung, "Learning common time-frequency-spatial patterns for motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 699–707, 2021, doi: 10.1109/TNSRE.2021.3071140.
[30] Y. Hou, T. Chen, X. Lun, and F. Wang, "A novel method for classification of multi-class motor imagery tasks based on feature fusion," *Neurosci. Res.*, vol. 176, pp. 40–48, Mar. 2022, doi: 10.1016/j.neures.2021.09.002.
[31] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, 2017.
[32] S. Chen, X. Zhang, L. Chen, and Z. Yang, "Automatic diagnosis of epileptic seizure in electroencephalography signals using nonlinear dynamics features," *IEEE Access*, vol. 7, pp. 61046–61056, 2019, doi: 10.1109/ACCESS.2019.2915610.
[33] J. Cao, L. Chen, D. Hu, F. Dong, T. Jiang, W. Gao, and F. Gao, "Unsupervised eye blink artifact detection from EEG with Gaussian mixture model," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 8, pp. 2895–2905, Aug. 2021, doi: 10.1109/JBHI.2021.3057891.
[34] M. Shahbakhti, M. Maugeon, M. Beiramvand, and V. Marozas, "Low complexity automatic stationary wavelet transform for elimination of eye blinks from EEG," *Brain Sci.*, vol. 9, no. 12, p. 352, Dec. 2019.
[35] A. K. Maddirala and K. C. Veluvolu, "ICA with CWT and k-means for eye-blink artifact removal from fewer channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1361–1373, 2022, doi: 10.1109/TNSRE.2022.3176575.

KUO-KAI SHYU (Senior Member, IEEE) received the B.S. degree in electrical engineering from the Tatung Institute of Technology, Taipei, Taiwan, in 1979, and the M.S. and Ph.D. degrees in electrical engineering from National Chung Kung University, Tainan, in 1984 and 1987, respectively.

In 1988, he joined National Central University, Taiwan, where he is currently a Professor with the Department of Electrical Engineering. From 1988 to 1999, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. He was the Chairperson of the Department, from 2004 to 2007. He is currently a Professor with the Pervasive Artificial Intelligence Research (PAIR) Laboratories, Taiwan. His teaching and research interests include variable structure control systems and signal processing with applications in motor control, power electronics, and biomedical systems.

SZU-CHI HUANG was born in Taichung, Taiwan, in 1999. He received the B.S. degree in electronic engineering from the National Taipei University of Technology, Taipei, Taiwan, in 2021, and the M.S. degree from National Central University, Taoyuan, Taiwan, in 2023, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include signal processing and machine learning with applications in biomedical systems.

KAI-JEN TUNG was born in Taichung, Taiwan, in 1998. He received the M.S. degree in electrical engineering from National Central University (NCU), Taoyuan, Taiwan, in 2022. His research interests include machine learning and biomedical signal analysis.

LUNG-HAO LEE (Member, IEEE) received the B.S. degree in statistics from National Taipei University, Taiwan, in 2003, the M.S. degree in information management from Yuan Ze University, Taoyuan, Taiwan, in 2005, and the Ph.D. degree in computer science and information engineering from National Taiwan University, in 2015. From 2015 to 2018, he was a Postdoctoral Fellow with National Taiwan Normal University. In 2018, he joined as an Assistant Professor with the Department of Electrical Engineering, National Central University, Taiwan. He is currently an Associate Professor. His research interests include natural language processing, information retrieval and extraction, biomedical and health informatics, and artificial intelligence technologies.

PO-LEI LEE (Member, IEEE) was born in 1973. He received the B.S. degree in electrical engineering from National Cheng Kung University, Taiwan, in 1995, and the Ph.D. degree from the Institute of Biomedical Engineering, National Yang Ming Chiao Tung University, Taiwan, in 2000. From 2001 to 2005, he was a Postdoctoral Fellow with the Taipei Veterans General Hospital, researching the signal and image-analysis procedures for electroencephalography and magnetoencephalography signals. In 2005, he joined the Department of Electrical Engineering, National Central University, Taiwan. He is currently a Professor with the Pervasive Artificial Intelligence Research (PAIR) Laboratories, Taiwan. His research interests include the signal and image processing of EEG and MEG signals and designing the EEG-based brain–computer interfaces.

YU-HAO CHEN received the B.S. degree in bio-industrial mechatronics engineering from National Chung Hsing University (NCHU), in 2020, and the M.S. degree from National Central University, Taoyuan, Taiwan, in 2023. His research interests include machine learning, brain–computer interface, and biomedical signal analysis.