Multimodal Fuzzy Fusion for Enhancing the Motor-Imagery-based Brain Computer Interface

Li-Wei Ko[*], Institute of Bioinformatics and Systems Biology and Center for Intelligent Drug Systems and Smart Bio-devices (IDS^{2}B), National Chiao-Tung University, Hsinchu, Taiwan

Yi-Chen Lu, Institute of Bioinformatics and Systems Biology and Center for Intelligent Drug Systems and Smart Bio-devices (IDS^{2}B), National Chiao-Tung University, Hsinchu, Taiwan

Humberto Bustince, Department of Statistics, Computer Science and Mathematics and Institute of Smart Cities, Public University of Navarre, Pamplona, Spain

Yu-Cheng Chang, Centre for Artificial Intelligence, CIBCI Lab, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia

Yang Chang, Institute of Bioinformatics and Systems Biology and Center for Intelligent Drug Systems and Smart Bio-devices (IDS^{2}B), National Chiao-Tung University, Hsinchu, Taiwan

Javier Fernandez, Department of Statistics, Computer Science and Mathematics and Institute of Smart Cities, Public University of Navarre, Pamplona, Spain

Yu-Kai Wang, Centre for Artificial Intelligence, CIBCI Lab, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia

Jose Antonio Sanz, Department of Statistics, Computer Science and Mathematics and Institute of Smart Cities, Public University of Navarre, Pamplona, Spain

Graçaliz Pereira Dimuro, Centro de Ciencias Computacionais, Universidade Federal do Rio Grande, Rio Grande, Brazil

Chin-Teng Lin[*], Centre for Artificial Intelligence, CIBCI Lab, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, Australia

ABSTRACT

Brain–computer interface technologies, such as steady-state visually evoked potential, P300, and motor imagery are methods of communication between the human brain and the external devices. Motor imagery–based brain–computer interfaces are popular because they avoid unnecessary external stimuli. Although feature extraction methods have been illustrated in several machine intelligent systems in motor imagery-based brain–computer interface studies, the performance remains unsatisfactory. There is increasing interest in the use of the fuzzy integrals, the Choquet and Sugeno integrals, that are appropriate for use in applications in which fusion of data must consider possible data interactions. To enhance the classification accuracy of brain-computer interfaces, we adopted fuzzy integrals, after employing the classification method of traditional brain–computer interfaces, to consider possible links between the data. Subsequently, we proposed a novel classification framework called the multimodal fuzzy fusion-based brain-computer interface system. Ten volunteers performed a motor imagery-based brain-computer interface experiment, and we acquired electroencephalography signals simultaneously. The multimodal fuzzy fusion-based brain-computer interface system enhanced performance compared with traditional brain–computer interface systems. Furthermore, when using the motor imagery-relevant electroencephalography frequency alpha and beta bands for the input features, the system achieved the highest accuracy, up to 78.81% and 78.45% with the Choquet and Sugeno integrals, respectively. Herein, we present a novel concept for enhancing brain–computer interface systems that adopts fuzzy integrals, especially in the fusion for classifying brain–computer interface commands.

Corresponding Author:
Chin-Teng Lin (Chin-Teng.Lin@uts.edu.au), Li-Wei Ko (lwko@mail.nctu.edu.tw).

I. INTRODUCTION

Brain-computer interfaces (BCIs) are a method of communication between the human brain and an external device [1]. Among existing BCIs, motor imagery (MI) has been popular in recent years as an alternative communication pathway, and it is based on an individual voluntarily modulating electroencephalography (EEG) signals. In comparison with current existing BCI technologies, such as steady-state visually evoked potential [2], [3] and event-related potential [4], MI-based BCIs have the advantage of avoiding unnecessary external stimulus as the commands for controlling the external devices. BCI users can individually imagine their movements to trigger the MI-based BCI systems, not by their actual actions. Furthermore, EEG dynamics reveal the related brain areas when behaviors are performed by an individual during MI-based BCI experiments [5]. Because of the characteristics of this asynchronous BCI system [6], a variety of feature extraction techniques have been applied to differentiate between the EEG dynamics of left- and right-hand imagination. One of the frequently used algorithms for feature extraction in MI-based BCI systems is the common spatial pattern (CSP). However, the variance of extracted features from two classes of MI-based EEG data distributed by the CSP means high BCI classification accuracy is difficult to achieve with only a single classifier. Despite feature extraction methods have been illustrated in several machine intelligent systems in MI-based BCI studies [4], [7], [8], the overall performance remains unsatisfactory because of inter-participant and intra-participant variability. Such variability severely affects the discrimination of the methods between left-hand and right-hand MI commands. A recent study demonstrated that EEG modulations can be trained through motor learning [9]; however, this is time-consuming, and the limitations of EEG pattern recognition remain. Other studies have applied deep-learning techniques to enhance the performance of MI-based BCI systems [10], [11]; nevertheless, an enormous EEG dataset is required for performing such deep-learning techniques [12]. Therefore, how to extract the appropriate features and enhance BCI classification accuracy are the major challenges for developing an MI-based BCI system, especially among the different varieties of dynamic EEG data. Consequently, novel algorithms must be adopted for enhancing BCI performance.

There is increasing interest in the use of fuzzy integrals [13]-[15], because of their wide applicability.

The most relevant examples of fuzzy integrals are the Choquet [16] and Sugeno integrals [14]. Both make use of a fuzzy measure to consider the relevance of possible coalitions (i.e., the possible links existing between data). This feature of fuzzy integrals makes them highly appropriate for applications in which fusion of data while considering their possible interactions is a relevant step, such as in cases of image processing [17], [18]; classification [19]-[22]; or decision making [23]. Some of the most widely used averaging functions, such as weighted means or the ordered weighted averaging (OWA) operators, are specific cases of fuzzy integrals (see [13]).

Furthermore, several generalizations of the classical notion of fuzzy integrals have been introduced in the literature, specifically relaxing linearity in the definition of the Choquet integral. Of note are the CF [21], [24] and CF1, F2 [19] integrals. These generalizations have proven successful in classification problems, where they are able to obtain results as effectively as state-of-the-art classifiers (see [19] and [21]). An advantage of these generalizations of the Choquet integral is that monotonicity is not required as in usual aggregation functions; however, in many cases, the Choquet integral is directionally monotone (that is, increasing along a fixed direction, see [25], [26]), which provides it with greater flexibility than other fuzzy integrals.

Considering the advantages, we adopted the fuzzy integrals after classification of the traditional BCI system, which can consider possible links for EEG signal processing, classification, and control commands. Our previous study [27] demonstrated the feasibility of implementing the fuzzy integral with particle swarm optimization to improve BCI performance. However, the proposed algorithm in reference [27] was only considering the varied features of different EEG frequencies, but not considering the features from different classifiers. In addition, EEG signal has time-variant and individual difference properties such that we intend to apply fuzzy integrals to solve the problem. Therefore, adopting fuzzy integrals could compensate for using different EEG frequencies and classifiers during classification in the present study. The proposed novel classification framework, which named as the **multimodal fuzzy fusion (MFF)-based BCI system**, is illustrated in Section II. The experimental results in Sections III and IV show that adopting fuzzy integrals into the BCI system can enhance BCI performance in comparison with traditional BCI systems. Furthermore, using the

MI-relevant EEG frequency alpha and beta bands as the input features fed into the proposed MFF-based BCI system maximized accuracy. Therefore, we present a novel concept of adopting fuzzy integrals in a BCI system, especially in the fusion for classifying BCI commands.

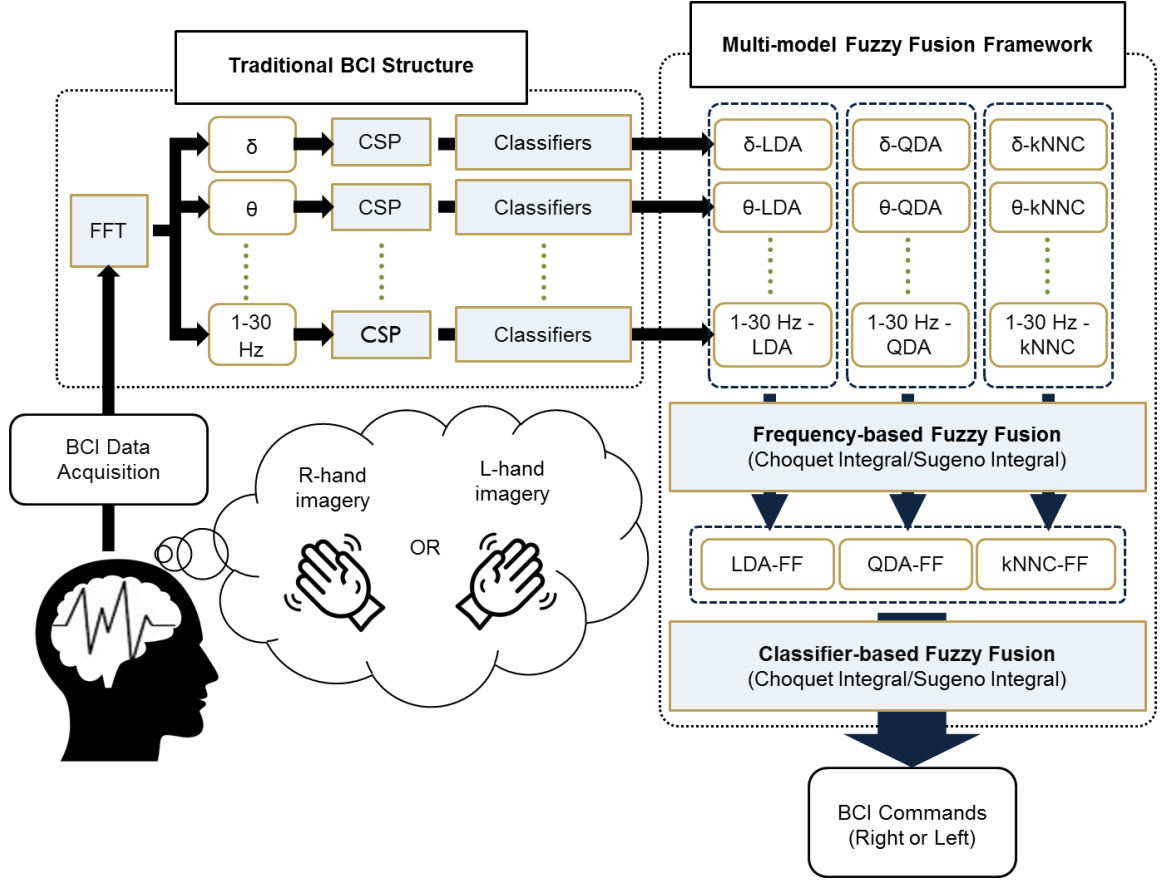## II. MULTIMODAL FUZZY FUSION-BASED BRAIN COMPUTER INTERFACE SYSTEM

Fig. 1. System architecture of the proposed MFF-based BCI system.

To enhance MI-based BCI performance, we proposed to apply a multimodal fuzzy fusion framework to an MI-based BCI system in this study. Fig. 1 shows the system architecture of the proposed BCI, which consisted primarily of two parts: the first part follows the traditional BCI structure, and the other part contains the MFF framework embedded into the MI-based BCI. The following sections detail the data analysis approaches.

A. Traditional BCI Structure

The traditional BCI system structure included four parts: (1) EEG data acquisition and preprocessing, (2) fast Fourier transform (FFT), (3) a feature extraction method utilizing the CSP, and (4) control command by classifiers. The first step in developing the traditional BCI system was acquiring the EEG data from the commercial EEG device and performing band-pass filtering and artifact removal on the collected EEG signals. The second step was EEG feature transformation and feature extraction. We

adopted FFT as the feature transformation to transfer the time-series EEG signals to different frequency bands, including the delta, theta, alpha, and beta bands. Subsequently, the CSP was used for feature extraction to extract the maximum spatial separability from the different EEG signals corresponding to the control commands. Last, pattern classification was performed on the extracted EEG signals using different classifiers to differentiate the commands for controlling the peripheral control device. Data analysis for each part is described as follows.

1.) *EEG Data Acquisition and Preprocessing*

Thirty-two EEG signals were collected by the Neuroscan system developed by Compumedics Ltd. (VIC, Australia). The sensor placements followed the standard international 10–20 system, and the reference channels were on the left and right mastoids. Contact impedances of all EEG electrodes must be below five kΩ. We selected four EEG channels placed at C3, C4, CP3, and CP4 to cover the motor cortex and sensorimotor cortex. After data acquisition, EEG signal preprocessing, including band-pass filtering with cutoff frequencies above 50 Hz and below 0.5 Hz, was conducted manually using the open source toolbox EEGLAB [28], [29]. The concept for the MI-based BCI experiment design is described in Section III.

2.) *FFT*

FFT was used in this study to reduce the complexity of discrete Fourier transform computation and to rapidly transform the EEG signals into different frequency components. FFT analysis transformed the time-series EEG signals in each channel into the frequency range from 1 to 30 Hz, covering the delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), and beta (14-30 Hz) bands using a 50-point moving window segment overlapping 25 data points.

3.) *CSP*

The CSP, which is a well-known mathematical procedure commonly used in EEG signal processing, was used in this study to transform multivariate EEG signals into well-separated subcomponents with maximum spatial variation [30], [31]. We adopted the CSP to transform the MI-based BCI datasets with two classes into a well-separated feature map for classifying BCI commands. Thus, the EEG feature vectors of four channels extracted through the CSP with

different frequency bands respectively in each EEG sample were estimated as the posterior probability in the classification step to evaluate MI-based BCI performance.

4.) *Classifications*

Three different classifiers, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and *k*-nearest neighbors classification (k-NNC), were adopted to classify the transformed MI-based BCI data. Using these classifiers, we observed which classification had optimal performance for MI-based BCI classification. LDA is a popular linear classification method widely applied in statistics and machine learning. Mean vectors and covariance matrices of distinct classes are the main parameters of LDA, which are calculated for searching the appropriate discrete features and separating them into two or more classes. QDA is a common multivariate classification similar to LDA. Unlike LDA, which employs a linear boundary between the data points of distinct classes, QDA separates the estimates of two or more classes with a quadric surface. The *k*-NNC is a nonparametric method for BCI classification. An unlabeled testing data point is classified by estimating the k neighbors (k=9 in this study) nearest to the testing data point among the training samples.

The classification accuracies of the MI-based BCI were calculated as the mean of the classification result for each sample. Statistical significance was estimated by paired *t*-test, with significance indicated by a *p*-value lower than 0.01 ($p < 0.01$). The posterior probabilities of left-hand and right-hand MI acquired from each classifier described previously were then fused using the Choquet or Sugeno fuzzy integral through the proposed MFF structure.

B. Multimodal Fuzzy Fusion Framework

To enhance the BCI performance, we proposed the MFF framework with the Choquet or Sugeno integrals to fuse the posterior probability obtained from the results of different classifiers. We considered that the EEG power changes of different frequencies affected the BCI performance more than using different classifiers did. Therefore, we implemented the frequency-based fuzzy fusion first before implementing the classifier-based fuzzy fusion to enhance the BCI performance for left-hand or

right-hand classification commands. The following sections define the fuzzy integrals adopted in this study.

1.) *Choquet Fuzzy Integral Fusion*

The Choquet integral is a generalization form of an integral using fuzzy (nonadditive) measures. In a discrete setting, if $g: 2^N \rightarrow [0,1]$ (where $N = \{1, \ldots, n\}$) is a fuzzy measure (a set-valued increasing function such that $g(\emptyset) = 0$ and $g(N) = 1$), the Choquet integral is an aggregate function $C_g: [0,1]^n \rightarrow [0,1]$ defined, for each $x = (x_1, \ldots, x_n) \in [0,1]^n$ and each membership function h, as:

$$C_g(h) = \sum_{i=1}^{n} \left(h(x_{(i)}) - h(x_{(i-1)}) \right) \cdot g(A_{(i)}), \qquad [M-C1] \qquad (1)$$

where $0 \leq h(x_{(1)}) \leq \ldots \leq h(x_{(n)})$, with the convention that $h(x_{(0)}) = 0$, and $A_{(i)} = \{(i), \ldots, (n)\}$ is the subset of indices of $n - i + 1$ largest components of $(h(x_1), \ldots, h(x_n))$.

As stated in the introduction, the weighted arithmetic mean and OWA operators are special cases of the Choquet integral, Choquet integral-based aggregation functions represent a larger class of aggregation functions. Recall that, given a vector $(w_1, \ldots, w_n) \in [0,1]^n$ with $w_1 + \cdots + w_n = 1$, the OWA operator associated to this vector is the function OWA: $[0,1]^n \rightarrow [0,1]^n$ defined by

$$\text{OWA}(x_1, \ldots, x_{(n)}) = w_1 x_{(1)} + \ldots + w_n x_{(n)}, \qquad (2)$$

where $x_{(1)} \geq \cdots \geq x_{(n)}$ is a decreasing reordering of the inputs $(x_1, \ldots, x_n)$ [8].

In order to generalize the Choquet integral to obtain a pre-aggregation function (i.e., a function with the same boundary conditions as an aggregation function but that is only increasing along some fixed direction; see [25], [26]), two distinct approaches (approach A and approach B) are considered as follows.

First, the approach A is to replace the product with a more general operation F [21]. In this sense, for a fuzzy measure $g: 2^N \rightarrow [0,1]$ and a function $F: [0,1]^2 \rightarrow [0,1]$, we define the C_F integral as the function $C_g^F: [0,1]^n \rightarrow [0,1]$ given by

$$C_g^F(h) = \sum_{i=1}^{n} F\left(h(x_{(i)}) - h(x_{(i-1)}), g(A_{(i)}) \right). \qquad (3)$$

If $F(x, y) = x \cdot y$, then we recover the standard Choquet integral as given by [M-C1]. For a general choice of F, the function C_g^F is not required to be an aggregation function as monotonicity can be violated. If the function F is such that either

(i) $F(0, x) = 0$ and $F(x, 1) = x$, or

(ii) for all $x \in [0, 1]$, $F(0, x) = 0$, $F(1, 1) = 1$, and F is $(1, 0)$-increasing (i.e.,

$\quad F(x', y) \geq F(x, y)$ if $x' > x$ for every $y \in [0, 1]$),

then, for any fuzzy measure g, the C_F integral defined by (3) is a pre-aggregation function (which is not generally an aggregation function). This is the case if we consider F as the Hamacher t-norm:

$$T_H(x, y) = \begin{cases} 0, & \text{if } x = y = 0 \\ \frac{xy}{x+y-xy}, & \text{otherwise} \end{cases} \tag{4}$$

Denote the resulting function as [M-C2].

Second, in the approach B, we are considering the distributivity of the product, the standard Choquet integral [M-C1] can be written as

$$C_g(h) = \sum_{i=1}^{n} \left(h(x_{(i)}) \cdot g(A_{(i)}) - h(x_{(i-1)}) \cdot g(A_{(i)}) \right). \qquad [M - C2] \quad (5)$$

If two general operations, F_1 and F_2 [19], [24], are such that

(i) F_1 is $(1, 0)$-increasing (i.e., $F(x, y) \geq F(z, y)$ whenever $x > z$),

(ii) $F_1(0, y) = F_2(0, y)$ for every $y \in [0,1]$,

(iii) $F_1(1,1) = 1$, and

(iv) $F_1 \geq F_2$,

then the resulting C_{F_1, F_2} integral is a pre-aggregation function. In particular, this is the case considering $F_1 = F_2 = min$; that is,

$$C_g^{min,min}(h) = \sum_{i=1}^{n} \min \left(h(x_{(i)}), g(A_{(i)}) \right) - \min \left(h(x_{(i-1)}), g(A_{(i)}) \right). \qquad [M - C3] \quad (6)$$

This function, contrary to the case of the usual Choquet integral, is not averaging; that is, its output is not required to be between the minimum and maximum of the inputs.

2.) *Sugeno Fuzzy Integral Fusion*

As aforementioned, another fruitful generalization of the usual Lebesgue integral is the Sugeno

integral [13], [14]. The Sugeno integral over the set $A = \{x_1, \ldots, x_n\}$ of a membership function h with respect to the fuzzy measure (confidence) g is defined as

$$S_g(h) = \sup_{\alpha \in [0,1]} [\min (\alpha, g(A \cap F_\alpha))], \qquad [M - S1] \qquad (7)$$

where $F_\alpha = \{ x \mid h(x) \geq \alpha \}$. Therefore, The Sugeno integral function can be descried as

$$S_g(x) = \max\{\min \left(h(x_{(i)}), g(A_{(i)}) \right) \mid i = 1, \ldots, n\}. \qquad (8)$$

With this definition, if $h(x) = x$ for every $x \in [0,1]$, S_g is an averaging aggregation function.

In general form, the S_F integral is the function S_g^F given by

$$S_g^F(h) = \max\{ F \left(h(x_{(i)}), g(A_{(i)}) \right) \mid i = 1, \ldots, n\}. \qquad (9)$$

If we consider $F(x, y) = T_H(x, y)$ (the Hamacher t-norm), that is,

$$S_g^{T_H}(h) = \max \{ T_H \left(h(x_{(i)}), g(A_{(i)}) \right) \mid i = 1, \ldots, n\}. \qquad [M - S2] \quad (10)$$

If we consider the pre-aggregation function (but not the aggregation function) $F(x, y) = BT(x, y) = x \cdot | 2y - 1|$, that is,

$$S_g^{BT}(h) = \max \{ h(x_{(i)}) \cdot |2g(A_{(i)}) - 1| \mid i = 1, \ldots, n\}. \qquad [M - S3] \quad (11)$$

III. MOTOR-IMAGERY-BASED BCI DATA COLLECTION AND ANALYSIS

In this study, ten participants aged 18-29 years without any neurological diseases were recruited in the MI-based BCI experiment. Before the experiment, the participants completed an informed consent form. The experimental protocol was approved through the Institutional Review Board of Taipei Veterans General Hospital (protocol number VGHUST102-G5-2-1). The participants sat in front of a monitor, and the instructions of the experimental protocol were shown to them before the experiment began. The participants performed the required tasks for the experiment and we acquired their EEG signals simultaneously. During the experiment, four EEG channels at C3, C4, CP3, and CP4 were selected for EEG data collection. The following sections describe the experimental procedure, BCI data analysis, and comparison of classification performance.
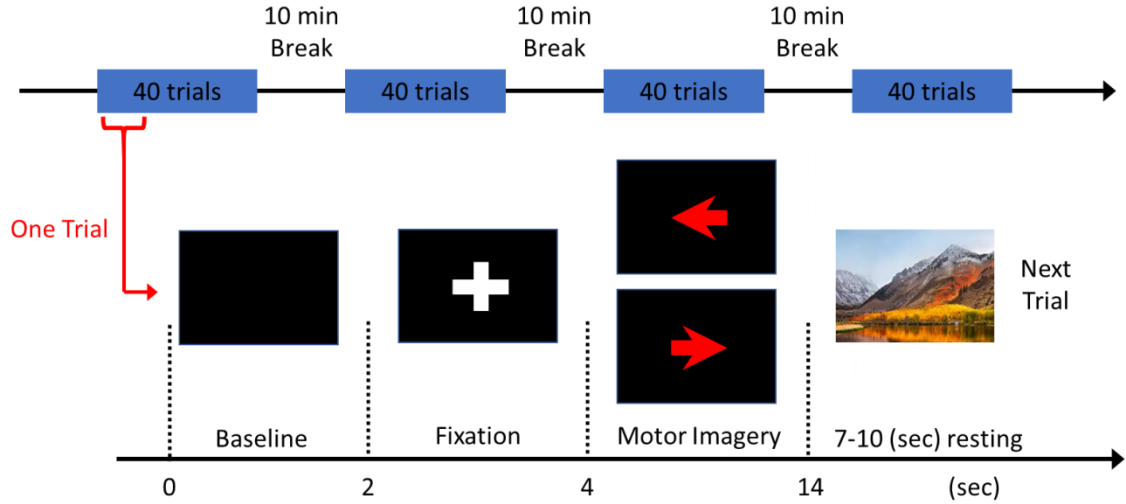
A. *Experimental Procedure*

Fig. 2. Flowchart of MI-based BCI experiment.

The experimental procedure is shown in Fig. 2. At the beginning of each trial, the screen was blank for two seconds (from 0 to 2 secs). Subsequently, a cross was displayed in the center of the screen for another two seconds (from 2 to 4 secs). When the cross disappeared, an arrow pointing either to the left or right randomly appeared on the screen for the following 10 seconds (from 4 to 14 secs). Upon seeing an arrow pointing to the left, the participant needed to imagine left-hand movement. Conversely, the participant needed to imagine right-hand movement when the arrow pointed to the right. After finishing

the movement imagery, a picture was displayed on the screen for a random duration (7–10 secs) in which the participant could rest before the next trial started. One trial was complete when the participant finished one movement imaginary activity; in total, there were four separate experiment sessions in this study. During the experiment, the participants were instructed to perform four sessions of the MI-based experiment task, and each session had 40 trials. A 10-min break period was provided between each session. EEG data were recorded with the time interval of the whole experiment according to the left-hand and right-hand MI-based experiment tasks as epochs from the C3, C4, CP3, and CP4 channels. Therefore, we collected 160 EEG samples totally consisting of four features and two labels to constitute the EEG datasets for evaluating the proposed MFF-based BCI system.

Table I shows the collected trials for the left-hand and right-hand movement imagery from each participant. We randomly generated ten datasets consisted of 80 training data points and 80 testing data points from the collected trials from each participant. Training data comprised 40 left-hand and 40 right-hand movement imaginary trials that were randomly selected from each participant's collection, and the remaining trials formed the testing data. We performed holdout validation for the proposed MFF–based BCI system. Because each participant had ten datasets, we averaged the accuracy from these datasets and investigated the stability of our proposed MFF-based BCI system. The following sections present the comparison of the classification performance of various BCI systems.

B. Comparison of the BCI performances between using fuzzy fusion and non-fuzzy fusion

To evaluate the classification performance of the proposed MFF-based BCI system, we first compared the experiment results between using fuzzy fusion and non-fuzzy fusion. Here, non-fuzzy fusion BCI was the traditional BCI structure, as shown in Fig. 2. The performances of non-fuzzy fusion classifiers, the Choquet and Sugeno integrals, are shown in Fig. 3. In non-fuzzy fusion classifiers, LDA, k-NNC, and QDA used five frequency bands, including four independent bands (i.e., delta, theta, alpha, and beta bands) and one full-band signal (1-30 Hz). In all three classifiers and all frequency bands, CSP projection was conducted to extract the EEG features.

TABLE 1 Collected left- and right-hand MI trials from each participant.

	LEFT-HAND MI TRIALS	RIGHT-HAND MI TRIALS
SUBJECT 1	85	75
SUBJECT 2	85	75
SUBJECT 3	82	78
SUBJECT 4	74	86
SUBJECT 5	74	86
SUBJECT 6	80	80
SUBJECT 7	80	80
SUBJECT 8	80	80
SUBJECT 9	82	78
SUBJECT 10	84	76

Before applying fuzzy fusion into the traditional BCI system, classification performances of using LDA classifier achieved 72.72% accuracy, and using QDA and k-NNC classifiers were below 70% accuracy. For the results of using the Choquet integral—one of the fuzzy fusion approaches—three cases, namely [M-C1], [M-C2], and [M-C3], were chosen to validate the BCI performance. For another fuzzy fusion approach, the Sugeno integral, three cases were chosen, [M-S1], [M-S2], and [M-S3], to validate the BCI performance too.

According to the fuzzy fusion mechanism of the Choquet and Sugeno integrals, the final decision was made by integrating classification results of different frequency bands. As shown in Fig. 3, using the Choquet [M-C1] and Sugeno [M-S3] integrals achieved the better classification accuracy of 75.93% and 76.11%, respectively, outperforming the traditional non-fuzzy fusion BCI system. Furthermore, we calculated the information transfer rate (ITR) [32] to evaluate the efficiency of the proposed MFF-based BCI system in this study. The index ITR is composed of the accuracy, computation time, and the number of classes during the classification process. In Table II, ITR in all case of fuzzy fusion [M-C1],

[M-C2], [M-C3], [M-S1], [M-S2], and [M-S3] were higher than the three non-fuzzy classifiers (i.e., LDA, k-NNK and QDC). Applying the case of [M-C1] and [M-S3] integrals could achieve the better ITR of 39.69 bits per minute and 40.46 bits per minute in Choquet and Sugeno integrals, respectively. Therefore, adopting fuzzy fusion could enhance the BCI performance of the system for classifying left-hand and right-hand EEG signals.
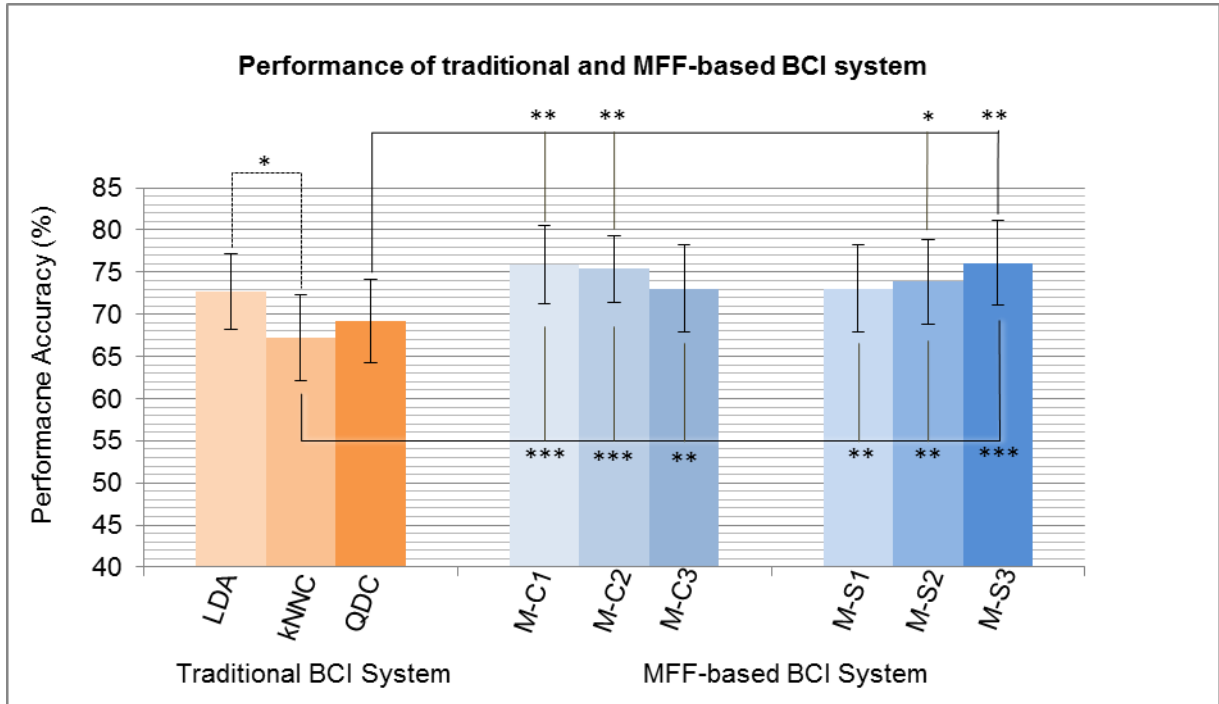

Fig. 3. Comparison of classification performance between traditional and MFF-based BCI system. The p-value of t-test shows the significant difference marked $p < 0.05$ as "*", $p < 0.01$ as "**", and $p < 0.001$ as "***".

TABLE 2 Information transfer rate of traditional and MFF-based BCI system

SYSTEM	CLASSIFICATION MODEL	ITR (BITS/MIN)
TRADITIONAL BCI	LDA	9.15
	k-NNC	4.84
	QDC	6.94
MFF-BASED BCI	CHOQUET [M-C1]	39.69
	CHOQUET [M-C2]	38.05

	CHOQUET [M-C3]	34.98
	SUGENO [M-S1]	33.84
	SUGENO [M-S2]	36.65
	SUGENO [M-S3]	40.46

As Fig. 3 shows, in every case, the considered fuzzy integrals provided significantly improved performance over that obtained from any of the three classifiers considered in our experiments. Furthermore, if we focus on the case of the Choquet integral and the considered extensions, we see that the standard [M-C1] and generalized [M-C2] Choquet integrals exhibited similar performance. This fact shows that consideration of extensions of the classical Choquet integral provides a promising method of tackling the fusion problem because the results can likely be improved through ad hoc choice of the fuzzy measure, but this was not the object of this study. This possibility is enhanced by the optimal performance of the generalized version of the Sugeno integral. In this case, the three cases of [M-S1], [M-S2], and [M-S3] not only outperformed the considered classifiers but also contrary to the case of the Choquet integral, were the best performers. Remarkably, case [M-S3] was significantly more effective than all other methods. However, this analysis is only a first approach to the BCI classification problem. We intend to develop a deeper analysis to optimize the choice of both measures and operators in the different extensions of the Choquet and Sugeno integrals. These rough approaches were sufficient to demonstrate that the introduction of nonlinearity (in the case of the Choquet integral) and of general operators instead of the minimum (in the case of the Sugeno integral) leads to a relevant improvement in the results.

C. Multimodal Fuzzy Fusion for Enhancing the BCI Performance

In Section III-B, we demonstrated that using fuzzy fusion achieved higher classification accuracy. Thus, we intended to fuse the classification results at different frequencies with the same classifier (called frequency-based fuzzy fusion, Fb-FF) and then fuse the results from different classifiers (called classifier-based fuzzy fusion, Cb-FF). Considering the simultaneous classification of different EEG

frequencies and classifiers, we called this the MFF framework, as shown in Fig. 2. Fig. 4 presents the comparison of BCI performance of the different fuzzy fusion approaches in the MFF-based BCI system using the Choquet (Fig. 4(a)) or Sugeno (Fig. 4(b)) integrals. We also compared the classification results obtained by fusing different EEG frequencies (Fb-FF) and by using a single frequency with different classifiers (Cb-FF). The mechanism of fuzzy fusion here was similar to the process of voting. Five features were applied to fuzzy fusion in the Fb-FF model for making the final decision; only three features were applied in the Cb-FF model.

Adopting the MFF framework exhibited superior performance in the cases [M-C1], [M-C2], and [M-C3] of the Choquet integrals and cases [M-S1], [M-S2], and [M-S3] of the Sugeno integrals when compared with Fb-FF and Cb-FF models. Moreover, the performances of MFF and Fb-FF models in the cases [M-C1], [M-C2], and [M-C3] of the Choquet integrals and the cases [M-S1], [M-S2], and [M-S3] of the Sugeno integrals showed statistically significant differences from the Cb-FF models.

As shown in Fig. 4(a) and 4(b), the proposed MFF-based BCI system achieved superior accuracy when adopting the Choquet integral (75.93%) and Sugeno integral (76.11%). These results indicated the proposed MFF framework could compensate for both posterior probabilities in classification using different frequencies and classifiers to yield improved decisions for BCI classification. Another finding was that the Fb-FF model outperformed the Cb-FF model. In comparison with the classification of the traditional BCI system (Fig. 3), adopting the Fb-FF model can provide better accuracy than the traditional BCI system can, but adopting the Cb-FF model cannot. Therefore, we inferred that the phenomena associated with the EEG dynamic changes at different frequencies of the neural mechanism play crucial roles in BCI classification. Consequently, we further compared the classification results of the EEG phenomena of the movement imaginary at different EEG frequencies; this is described in Section III-D.
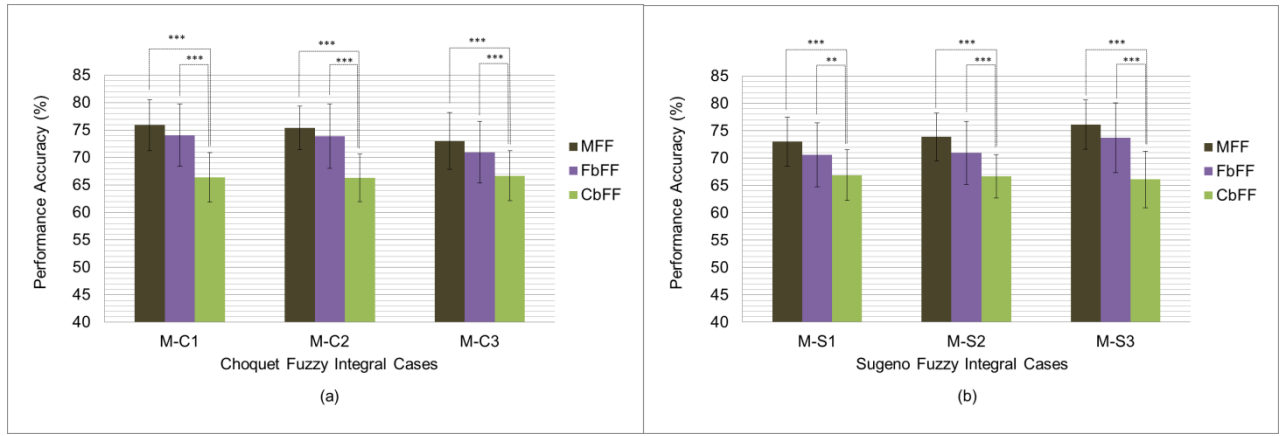
Fig. 4. Comparison of different fuzzy fusion approaches. (A) Choquet fuzzy integral by M-C1, M-C2, and M-C3. (B) Sugeno fuzzy integral by M-S1, M-S2, and M-S3. The *p*-value of *t*-test shows the significant difference marked $p < 0.05$ as "*", $p < 0.01$ as "**", and $p < 0.001$ as "***".

D. Motor imagery related EEG features selected in the multimodal fuzzy fusion BCI system

For real-world BCI applications, the performance sometimes trades off against efficiency (i.e., higher accuracy at the cost of fewer features). Selecting the appropriate EEG features for BCI classification is a major challenge. According to past MI-based BCI studies [6]-[11], the alpha and beta bands in the motor area are desynchronized while performing movement imagination tasks. Therefore, using all bands, the alpha band, the beta band, and both the alpha and the beta bands were considered as candidate EEG features for testing the MFF-based BCI system. Fig. 5 compares the classification performance of the MFF-based BCI system using those selected EEG features. Fig. 5(a) and 5(b) present the results of using the Choquet and Sugeno integrals, respectively. Using the alpha and the beta bands for the EEG features exhibited optimal performance among all conditions using either the Choquet or Sugeno integrals. Furthermore, the classification when adopting the Choquet integral in the alpha and the beta bands exhibited statistically significant differences from those in the beta band in the cases [M-C1] and [M-C2] and from those in all bands and the beta band in case [M-C3]. The classification results using the Sugeno integrals are shown in Fig. 5(b). Using the alpha band showed statistically significant differences from all other conditions in case [M-S1]. Using the alpha and the beta bands showed statistically significant differences from those in the beta band in cases [M-S2] and [M-S3].

Based on the experiment results and referring to past MI-based BCI studies [6]-[11], [27], the alpha band was regarded as the optimal feature for efficiency-performance balance. These classification

results were also consistent with past EEG findings for MI tasks. Feeding the alpha and the beta bands as the optimal features into the proposed MFF-based BCI system exhibited the best classification performance (approximately 80%) regardless of whether using the Choquet or Sugeno integrals. By contrast, the classification performance using only the beta had a lower accuracy (approximately 65%); however, with both the alpha and beta bands used in the proposed MFF BCI system, the classification accuracy increased to nearly 80%. Comparison of the classification accuracy is shown in Fig. 5; this finding is noteworthy because the proposed MFF-based BCI system optimized performance by accessing only one feature (i.e., the alpha band). Based on these findings, the real-world application of the proposed MFF-based BCI system is feasible.
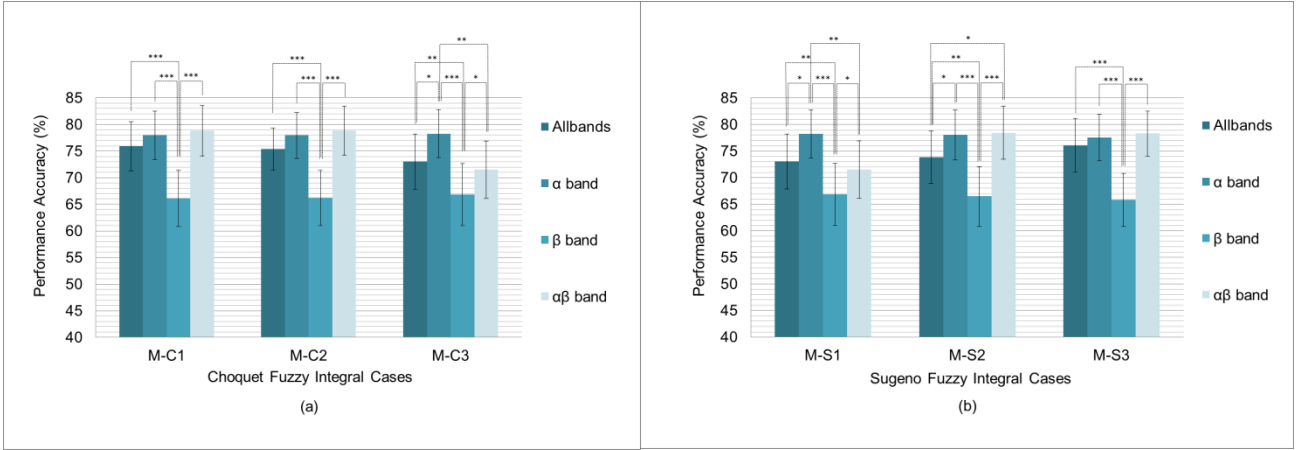

Fig. 5. Comparison of different feature selections. (A) Choquet fuzzy integral by M-C1, M-C2, and M-C3. (B) Sugeno fuzzy integral by M-S1, M-S2, and M-S3. The p-value of t-test shows the significant difference marked $p < 0.05$ as "*", $p < 0.01$ as "**", and $p < 0.001$ as "***".

IV. MFF-BASED BCI SYSTEM PERFORMANCE EVALUATION ON THE BCI COMPETITION DATASET

A. Benchmark Datasets

Considering the generalizability and the feasibility of the proposed MFF-based BCI system, we adopt the benchmark BCI datasets (i.e., datasets of BCI Competition IV [33]) for evaluating the BCI performance in this study. The datasets were collected from the cue-based BCI experimental paradigm including four-class (left hand, right hand, foot, and tongue) imaginary task as shown in [33] using

22-channel EEG signals from nine volunteer participants. Two sessions of imagery tasks were recorded, each session consisted of six runs; each run was comprised of 48 trials, yielding totally 288 trials per session. Trials from four classes were evenly distributed, which meant each class had 72 trials.

In order to evaluate the proposed MFF–based BCI system, we analyzed four channels out of the 22 channels (i.e., channel number 8, 12, 14, 18 of benchmark dataset) EEG data which were associated with the motor area and selected left- and right-hand imaginary trials of EEG data. We performed hold-out validation as well as section III-A by randomly generating ten datasets from each participant to estimate the stability of our proposed MFF–based BCI system. There were 72 training data samples and 72 testing samples consisted of five features, the delta, theta, alpha, and beta bands and the average frequency power from 1 to 30 Hz in each dataset. The CSP filter was calculated with four columns, which were, contained the second and second last columns. The classifiers (i.e., LAD, k-NNC, and QDA) were applied to estimate the classification performance and the accuracy from the means of ten datasets of each participant was averaged to perform the proposed MFF–based BCI system evaluation. Following sections showed the classification of benchmark datasets, and the evaluation of the proposed MFF–based BCI system performance.

B. *Evaluation of the BCI performances with fuzzy fusion and non-fuzzy fusion*

As the same comparison process in Section III-B, the classification performances of using Choquet or Sugeno integrals were shown in Fig. 6. Without applying fuzzy fusion, classification performances of all classifiers were below 60% accuracy. The CSP4 method proposed in the reference [33] in which 22 channel EEG signals were collected achieved 65.20% accuracy. As shown in Fig. 6, the proposed MFF-based BCI system of using the Choquet [M-C1] and Sugeno [M-S3] integrals and only four-channel EEG signals achieved the classification accuracy of 64.06% and 63.36%, respectively, demonstrating that the proposed fuzzy fusion framework outperformed three non-fuzzy classifiers in accuracy and classification performance equivalent with the BCI system in reference [33].

However, our proposed MFF-based BCI system of using only four channel EEG signals is more practical and feasible to realize the BCI applications in the real-world environment. Moreover, our

proposed MFF BCI system enhanced the performance with the data not only collected from our motor imagery experiment but also applied from the benchmark datasets revealed that it is potential to perform 2-class condition and more BCI applications.
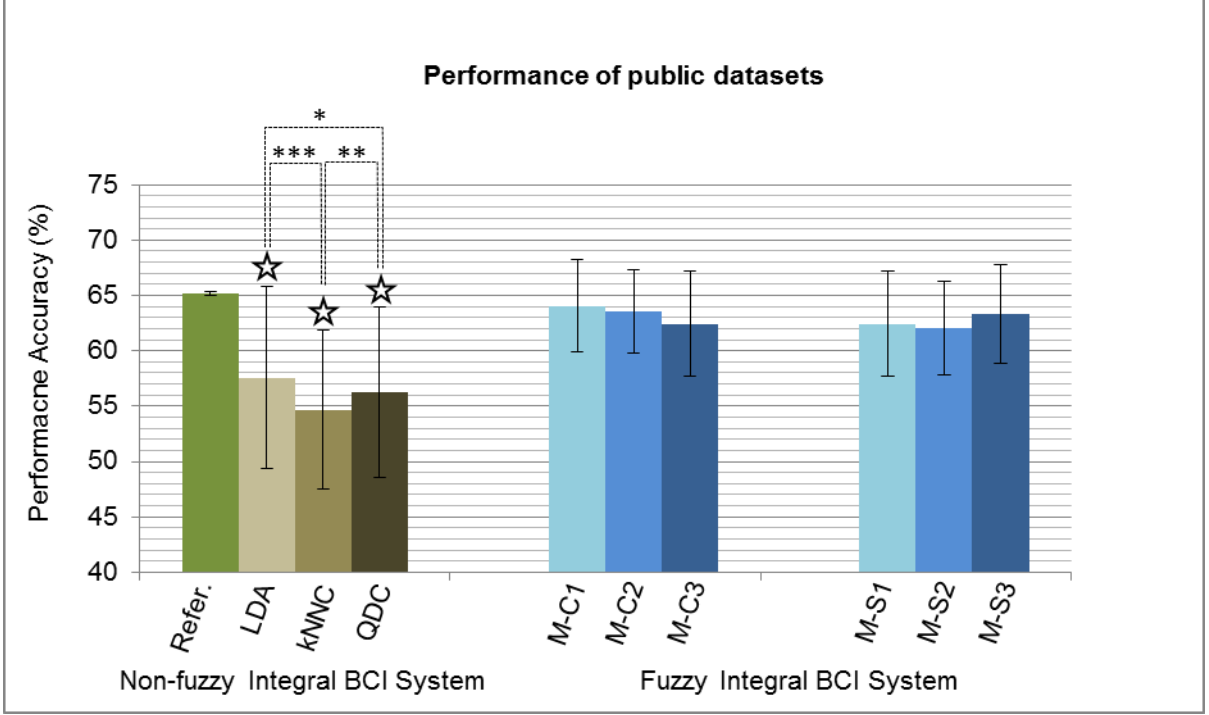

Fig. 6. Evaluation of BCI system performance between using fuzzy and nonfuzzy integrals. The p-value of t-test shows the significant difference marked $p < 0.05$ as "*", $p < 0.01$ as "**", and $p < 0.001$ as "***". The marker of white star means the significant difference (p-value < 0.001) between all case of fuzzy fusion and each non-fuzzy classifier.

C. Evaluation of Multimodal Fuzzy Fusion Framework

As the same comparison process in Section III-C, we further investigate the MFF framework and the other different fuzzy fusion approaches such as Fb-FF and Cb-FF models. The proposed MFF–based BCI system achieved superior accuracy when adopting the Choquet integral (64.06%) or Sugeno integral (63.36%) on the BCI competition datasets (Fig. 7(a) and 7(b)). In the case of comparison with the traditional BCI system (Fig. 6), adopting the Fb-FF and the Cb-FF model both can provide better accuracy than the traditional BCI system on the BCI competition datasets. These results indicated once again that fuzzy fusion could enhance the classification performance in accuracy. Another finding was the Fb-FF model outperformed the Cb-FF model that demonstrated that different frequencies features

played the crucial roles in BCI classification. This evidence is shown as the same as in the section III-C (Fig. 4). Consequently, we demonstrated that the proposed MFF outperform the optimal classification performance either testing on the BCI competition datasets or using our BCI motor imaginary datasets for classifying left-hand and right-hand EEG signals.
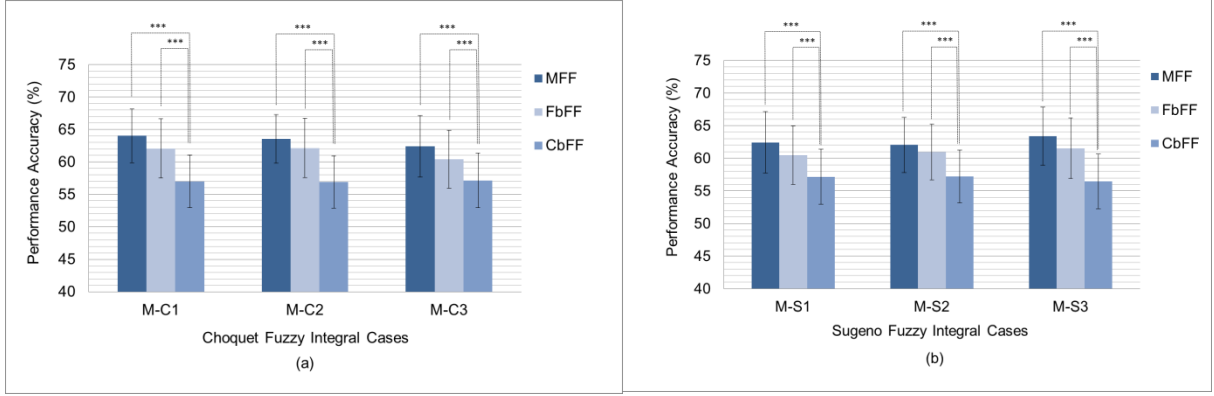

Fig. 7. Comparison of different fuzzy fusion approaches on the BCI competition datasets. (A) Choquet fuzzy integrals by M-C1, M-C2, and M-C3. (B) Sugeno fuzzy integrals by M-S1, M-S2, and M-S3. The p-value of t-test shows the significant difference marked $p < 0.05$ as "*", $p < 0.01$ as "**", and $p < 0.001$ as "***".

In this study, we illustrated a two-class classification problem using the proposed MFF-based BCI system. Considering the practical applications in the real-world environment, two-class or multi-classes MI-based BCI systems were comprehensively adopted into many applications such as controlling a wheelchair, unmanned aerial vehicle, and robotic apparatus for stroke rehabilitation [34]-[37]. In the study of 4-class MI-based BCI [38], the accuracy exhibited a significant reduction compared with 2-class MI-based BCI system. Therefore, we trust that the 2-class MI-based BCI system is more feasible and has to promote to human well-being friendly. The major challenges of the MI-based BCI system development were to generate the discriminative EEG features from the individual variation of different participants. Motor learning [11] can improve the training of MI-based EEG signals, but we still need the computational intelligent techniques like the proposed fuzzy fusion methods to enhance the BCI performance [27], [39].

V. CONCLUSION

In this study, we proposed a novel classification framework called the MFF-based BCI system. Two famous fuzzy integrals named the Choquet and Sugeno integrals were implemented into the MFF framework after the classification structure of the traditional BCI system to enhance the classification performance. The novelty of the proposed MFF framework was to fuse the posterior probabilities obtained from both classification results when using different frequencies and classifiers. Fuzzy integrals accounted for the possible relationships between different classification results. Adopting fuzzy integrals into the BCI system improved BCI performance over the traditional BCI system. Moreover, selecting the alpha and the beta bands in the motor area feeding into the proposed MFF-based BCI system achieved the highest accuracy: 78.81% and 78.45% when using the Choquet and Sugeno integrals, respectively. Therefore, adopting fuzzy integrals in BCI systems could provide a novel method of developing highly accurate BCI systems.

REFERENCES

[1] L. W. Ko, S. K. R. Singanamalla, O. Komarov, and C. C. Chen, "Development of single channel hybrid BCI system using motor imagery and SSVEP," *Journal of Healthcare Engineering*, vol. 2017:3789386, Aug. 2017.

[2] Y. C. Lee, W. C. Lin, F.-Y. Cherng, and L. W. Ko, "A visual attention monitor based on steady-state visual evoked potential," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 399-408, March 2016.

[3] Y. Li and C. S. Nam, "Collaborative brain-computer interface for people with motor disabilities," *IEEE Computational Intelligence Magazine*, vol. 11, no. 3, pp. 56-66, August 2016.

[4] Y. K. Wang, T. P. Jung, and C. T. Lin, "Theta and alpha oscillations in attentional interaction during distracted driving," *Frontiers in Behavioral Neuroscience*, vol. 12, Feb. 2018.

[5] M. Jeannerod, "Neural simulation of action: A unifying mechanism for motor cognition," *Neuroimage*, vol. 14, no. 1, pp. S103-S109, 2001.

[6] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 2, pp. 258-265, 2004.

[7] V. Jayaram, M. Alamgir, Y. Altun, B. Schölkopf, and M. Grosse-Wentrup, "Transfer learning in brain-computer interfaces," *IEEE Computational Intelligence Magazine*, vol. 11, iss. 1, pp 20-31, February 2016.

[8] S. Saeedi, R. Chavarriaga, R. Leeb, and J. R. Millan, "Adaptive assistance for brain-computer interface by online prediction of command reliability," *IEEE Computational Intelligence Magazine*, vol. 11, iss. 1, pp 32-39, February 2016.

[9] R. Ge, H. Zhang, L. Yao, and Z. Long, "Motor imagery learning induced changes in functional connectivity of the default mode network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 1, pp. 138-148, 2015.

[10] D. Zhang, L. Yao, X. Zhang, S. Wang, W. Chen, and R. Boots, "EEG-based intention recognition from spatio-temporal representations via cascade and parallel convolutional recurrent neural networks," *arXiv preprint arXiv:1708.06578*, 2017.

[11] S. Kumar, A. Sharma, K. Mamun, and T. Tsunoda, "A deep learning approach for motor imagery EEG signal classification," in *Proc. of 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, pp. 34-39, 2016.

[12] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Computational Intelligence Magazine*, vol. 13, no. 3, pp 55-75, August 2018.

[13] G. Beliakov, H. B. Sola, and T. C. Sánchez, *A Practical Guide to Averaging Functions*. Springer, 2016.

[14] M. Sugeno, "Theory of fuzzy integrals and its applications," *Doctoral Thesis, Tokyo Institute of Technology*, 1974.

[15] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Computational Intelligence Magazine*, vol. 11, no. 1, pp. 41-53, February 2016.

[16] G. Choquet, "Theory of capacities," in *Annales de l'institut Fourier*, 1954, vol. 5, pp. 131-295.

[17] H. Bustince, E. Barrenechea, M. Sesma-Sara, J. Lafuente, G. P. Dimuro, R. Mesiar, and A. Kolesárová, "Ordered directionally monotone functions. Justification and application," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 4, Aug. 2018.

[18] D. Paternain, M. J. Campión, R. Mesiar, I. Perfilieva, and H. Bustince, "Internal fusion functions," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 2, pp. 487-503, 2018.

[19] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, M. J. Asiain, M. Elkano, and H. Bustince, "CC-integrals: Choquet-like copula-based aggregation functions and its application in fuzzy rule-based classification systems," *Knowledge-Based Systems*, vol. 119, pp. 32-43, 2017.

[20] D. Paternain, H. Bustince, M. Pagola, P. Sussner, A. Kolesárová, and R. Mesiar, "Capacities and overlap indexes with an application in fuzzy rule-based classification systems," *Fuzzy Sets and Systems*, vol. 305, pp. 70-94, 2016.

[21] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, H. Bustince, and R. Mesiar, "CF-Integrals: A new family of pre-aggregation functions with application to fuzzy rule-based classification systems," *Information Sciences*, vol. 435, pp. 94-110, April 2018.

[22] C. Wu, T. Yoshinaga, X. Chen, L. Zhang, and Y. Ji, "Cluster-based content distribution integrating LTE and IEEE 802.11p with fuzzy logic and Q-learning," *IEEE Computational Intelligence Magazine*, vol. 13, no. 1, pp. 41-50, February 2018.

[23] H. Bustince, M. Galar, B. Bedregal, A. Kolesarova, and R. Mesiar, "A new approach to interval-valued Choquet integrals and the problem of ordering in interval-valued fuzzy set applications," *IEEE Transactions on Fuzzy systems*, vol. 21, no. 6, pp. 1150-1162, 2013.

[24] R. Mesiar, A. Kolesárová, H. Bustince, G. P. Dimuro, and B. Bedregal, "Fusion functions based discrete Choquet-like integrals," *European Journal of Operational Research*, vol. 252, no. 2, pp. 601-609, 2016.

[25] G. Lucca, J. A. Sanz, G. P. Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, and H. Bustinc, "Preaggregation functions: construction and an application," *IEEE Transactions on Fuzzy Systems*, vol. 24, no. 2, pp. 260-272, 2016.

[26] H. Bustince, J. Fernández, A. Kolesárová, and R. Mesiar, "Directional monotonicity of fusion functions," *European Journal of Operational Research*, vol. 244, no. 1, pp. 300-308, 2015.

[27] S. L. Wu, Y. T. Liu, T. Y. Hsieh, Y. Y. Lin, C. Y. Chen, C. H. Chuang, and C. T. Lin, "Fuzzy integral with particle swarm optimization for a motor-imagery-based brain–computer interface," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 1, pp. 21-28, 2017.

[28] C. Brunner, A. Delorme, and S. Makeig, "EEGLAB – An open source Matlab toolbox for electrophysiological research," *Biomedizinische Technik (Berl)*, 2013.

[29] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9-21, 2004.

[30] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41-56, 2008.

[31] C. Guger, H. Ramoser, and G. Pfurtscheller, "Real-time EEG analysis with subject-specific spatial patterns for a braincomputer interface (BCI)," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 447–456, 2000.

[32] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "EEG-based communication: Improved accuracy by response verification," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 3, pp. 326-333, 1998.

[33] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller, "Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis," *Pattern Recognition Letters*, vol. 28, no. 8, pp. 957-964, 2007.

[34] R. Ron-Angevin, F. Velasco-Álvarez, Á. Fernández-Rodríguez, A. Díaz-Estrella, M. J.

Blanca-Mena, and F. J. Vizcaíno-Martín, "Brain-computer interface application: Auditory serial interface to control a two-class motor-imagery-based wheelchair," *Journal of NeuroEngineering and Rehabilitation*, vol. 14, no. 1, p. 49, May 2017.

[35] K. Ang et al., Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback. 2010, pp. 5549-5552.

[36] T. Shi, H. Wang, and C. Zhang, "Brain computer interface system based on indoor semi-autonomous navigation and motor imagery for unmanned aerial vehicle control," *Expert Systems with Applications*, vol. 42, no. 9, pp. 4196-4206, 2015.

[37] C. T. Lin, C. Y. Chiu, A. K. Singh, J. T. King, and Y. K Wang, "A wireless multifunctional SSVEP-based brain computer interface assistive system," *IEEE Transactions on Cognitive and Developmental Systems*, 2018.

[38] T. Li, J. Zhang, T. Xue, and B. Wang, "Development of a novel motor imagery control technique and application in a gaming environment," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 5863512, 16 pages, 2017.

[39] Y. K. Wang, T. P. Jung, and C. T. Lin, "EEG-based attention tracking during distracted driving," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 6, pp. 1085-1094, Nov. 2015.