

Classification of Motor-Imagery Tasks Using a Large EEG Dataset by Fusing Classifiers Learning on Wavelet-Scattering Features

Tuan D. Pham¹, *Senior Member, IEEE*

Abstract—Brain-computer or brain-machine interface technology allows humans to control machines using their thoughts via brain signals. In particular, these interfaces can assist people with neurological diseases for speech understanding or physical disabilities for operating devices such as wheelchairs. Motor-imagery tasks play a basic role in brain-computer interfaces. This study introduces an approach for classifying motor-imagery tasks in a brain-computer interface environment, which remains a challenge for rehabilitation technology using electroencephalogram sensors. Methods used and developed for addressing the classification include wavelet time and image scattering networks, fuzzy recurrence plots, support vector machines, and classifier fusion. The rationale for combining outputs from two classifiers learning on wavelet-time and wavelet-image scattering features of brain signals, respectively, is that they are complementary and can be effectively fused using a novel fuzzy rule-based system. A large-scale challenging electroencephalogram dataset of motor imagery-based brain-computer interface was used to test the efficacy of the proposed approach. Experimental results obtained from within-session classification show the potential application of the new model that achieves an improvement of 7% in classification accuracy over the best existing classifier using state-of-the-art artificial intelligence (76% versus 69%, respectively). For the cross-session experiment, which imposes a more challenging and practical classification task, the proposed fusion model improves the accuracy by 11% (54% versus 65%). The technical novelty presented herein and its further exploration are promising for developing a reliable sensor-based intervention for assisting people with neurodisability to improve their quality of life.

Index Terms—Brain-computer interface, motor imagery, electroencephalogram, rehabilitation, wavelet scattering, fuzzy recurrence plots, support vector machines, weak classifier fusion.

Manuscript received 1 October 2022; revised 22 January 2023; accepted 29 January 2023. Date of publication 31 January 2023; date of current version 6 February 2023.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Shanghai Second Rehabilitation Hospital Ethics Committee under Approval No. ECSHSRH 2018-0101, and performed in line with the Helsinki Declaration.

The author is with the Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Al Khobar 31952, Saudi Arabia (e-mail: tdpham123@gmail.com).

Digital Object Identifier 10.1109/TNSRE.2023.3241241

I. INTRODUCTION

MOTOR imagery (MI) is a cognitive process during which a human participant is asked to imagine performing a movement without actually doing it and exercising muscular tension [1]. In other words, MI requires the neural activation of the brain involving in the preparation and execution of an action accompanied with a voluntary restraining of the actual task [2]. MI-based brain-computer interface (MI-BCI) is a signal-based coordination that enables the control or recognition of brain activity over computer-based devices [3], [4]. In comparison with other BCI methods, MI-BCI allows direct interaction between a user and external devices without limb movement or stimulation [5].

MI-BCI technology has been realized as a new intervention for control and communication with disabled patients who have severe motor disorders [6], [7], difficulty in speech expression [8], and rehabilitation of people with neurological disease [9] or post-stroke patients [10], [11], [12]. Although the technology shows promising results and has been increasingly investigated, the delivery of a reliable MI-BCI system for real-world applications is still far from reality, and still remains a technical challenge for research and development [13], [14].

Machine-learning algorithms have been developed for studying BCI. A decision-tree approach was introduced for choosing a BCI device for patients who are not cognitively impaired but have movement or communication disability [15]. Signal processing for feature extraction of MI from electroencephalographic (EEG) data in BCI and different types of pattern classifiers have been investigated for predicting mental intentions [16]. Some most recently developed methods for classifying MI-BCI tasks recorded with EEG signals include convolutional neural networks (CNNs) [17], [18], [19] and pretrained deep-learning models [20], which are considered state-of-the-art artificial intelligence (AI) models for pattern classification, methods of truncation thresholds and empirical mode decomposition for feature extraction of EEG signals with statistical learning methods [21], and the use of fewer EEG channels with the Dempster-Shafer theory of evidence for decision making under uncertainty [22].

In the field of complex signal analysis, wavelet scattering transforms or networks [23], [24], [25] offer the extraction of low-variance coefficients from time series (being robust to

translation) and images (being robust to both translation and rotation), which can be useful for machine learning and classification. Particularly for the analysis of EEG signals, wavelet scattering networks have recently been used for extracting differential features of the data for classifying alcohol-affected and healthy subjects [26], recognizing emotional states [27], and detecting different types of heartbeats [28].

Being motivated by the usefulness of wavelet scattering transforms for discovering discriminative features in 1-D and 2-D data, this study attempts to utilize the wavelet scattering in two folds. First, wavelet time scattering is adopted for extracting features directly from MI-BCI EEG signals. Second, the EEG signals are transformed into texture images whose features are then extracted using wavelet image scattering, which was reported to be capable of capturing compact information about the spatial arrangement and pixel intensities of texture [29]. The transformation of 1-D signals into 2-D images of texture is carried out by the method of fuzzy recurrence plots (FRPs) [30] developed for nonlinear data analysis of recurrence in chaos. Texture at both small and large scales can be visualized in an FRP to gain insight into the temporal behavior of the reconstructed phase space of a dynamical system. Next, wavelet time and image scattering coefficients are used separately to train two classifiers for differentiating MI-BCI tasks, which are expected to produce complementary results based on their learning on different features of the EEG signals. Finally, a rule-based decision is developed for combining the outputs obtained from the two classifiers to result in the final classification.

The rest of this paper is organized as follows. Section II describes a big challenging EEG database of MI-BCI used in this study. Section III presents technical methods utilized for feature extraction of EEG signals included in the database and decision rules for combining outputs from complementary weak classifiers. Section IV shows experiments of within-session and cross-session classifications, including results in terms of several statistical measures, and provides comparisons between the proposed approach and different models as well as discussion. Finally, Section V includes remarks of the finding, limitation of the study, and suggestions of remaining issues for future research.

II. MI-BCI EEG DATA

The publicly available big MI-BCI EEG database [5] used in this study consists of EEG recordings from 25 participants (males = 13, females = 12, and years of age = 20-24). Each participant performed 5 independent sessions over 2 or 3 days, where each session included 100 trials of left-hand (class 1) and right-hand (class 2) grasping MI, and each trial lasted about 7.5 sec. The study on this human data acquisition was approved by the Shanghai Second Rehabilitation Hospital Ethics Committee (approval number: ECSHSRH 2018-0101) and was conducted in accordance with the Helsinki Declaration.

The participants were tasked to imagine the movements of the left-hand and right-hand actions of grasping according to the video and audio cues. The kinetic MI tasks were recorded with a distribution of 32 EEG cap electrodes, where, during

the data collection, the electrode impedance $< 20 \text{ k}\Omega$, and the sampling frequency = 250 Hz.

To be readily used for classification by machine learning, baseline fluctuations were removed from all the data, and the signals were 0.5–40 Hz band-pass filtered with a finite impulse response. Because of the deletion of some degraded segments, the EEG data in some trials carried out by some participants were missing. In general, the whole EEG database are presented in a tensor form as: 25 (subjects) \times 5 (sessions) \times 100 (trials) \times 32 (channels) \times 1,000 (samples), constituting approximately 400,000 EEG signals of 1,000 samples in length. This EEG database can be freely downloaded at figshare [31].

Figure 1 shows the EEG signals recorded from two subjects performing MI trials by imagining the left-hand (class 1) or right-hand (class 2) grasping.

III. METHODS

A. Transformation of EEG Signals Into Texture Images

Let $\mathbf{u} = (u_1, u_2, \dots, u_n)$ be an ECG signal. The phase-space reconstruction of the signal, denoted as \mathbf{Y} , can be obtained using a delay embedding theorem in chaos and nonlinear dynamics as [32]

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N), \quad (1)$$

with

$$\mathbf{y}_k = [u_k, u_{k+\tau}, \dots, u_{k+(m-1)\tau}], k = 1, \dots, N, \quad (2)$$

where m and τ are values for embedding dimension and time delay, respectively, and $N = n - (m - 1)\tau$, where n is the length of an EEG signal \mathbf{u} previously expressed.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\}$ be a set of c cluster centers of \mathbf{Y} , and $\mu(a, b) \in [0, 1]$ be a fuzzy membership grade expressing the degree of similarity between elements a and b (a higher degree indicates stronger similarity between the two elements). Using the reconstructed phase space of \mathbf{u} , an $N \times N$ texture image of the EEG signal can be generated in terms of an FRP with the following steps [30].

- 1) Impose the reflexive property for an FRP as

$$\mu(\mathbf{y}_k, \mathbf{y}_k) = 1, k = 1, \dots, N. \quad (3)$$

- 2) An FRP has the symmetric property as

$$\mu(\mathbf{y}_k, \mathbf{v}_l) = \mu(\mathbf{v}_l, \mathbf{y}_k), k = 1, \dots, N, l = 1, \dots, c. \quad (4)$$

- 3) Deduce the transitive property for an FRP by

$$\mu(\mathbf{y}_k, \mathbf{y}_j) = \max [\min \{\mu(\mathbf{y}_k, \mathbf{v}_l), \mu(\mathbf{v}_l, \mathbf{y}_j)\}], \\ l = 1, \dots, c; k \neq j. \quad (5)$$

- 4) Estimate $\mu(\mathbf{y}, \mathbf{v})$ for computing the inference expressed in Equation (5).

- 5) Finally, an $N \times N$ texture image \mathbf{I} , which is an FRP, of the EEG signal \mathbf{u} can be constructed as

$$\mathbf{I} = [\mu(\mathbf{y}_k, \mathbf{y}_j)], k, j = 1, \dots, N. \quad (6)$$

The estimate of $\mu(\mathbf{y}, \mathbf{v})$ can be obtained using the fuzzy c -means (FCM) algorithm [33] as follows. Given a number of

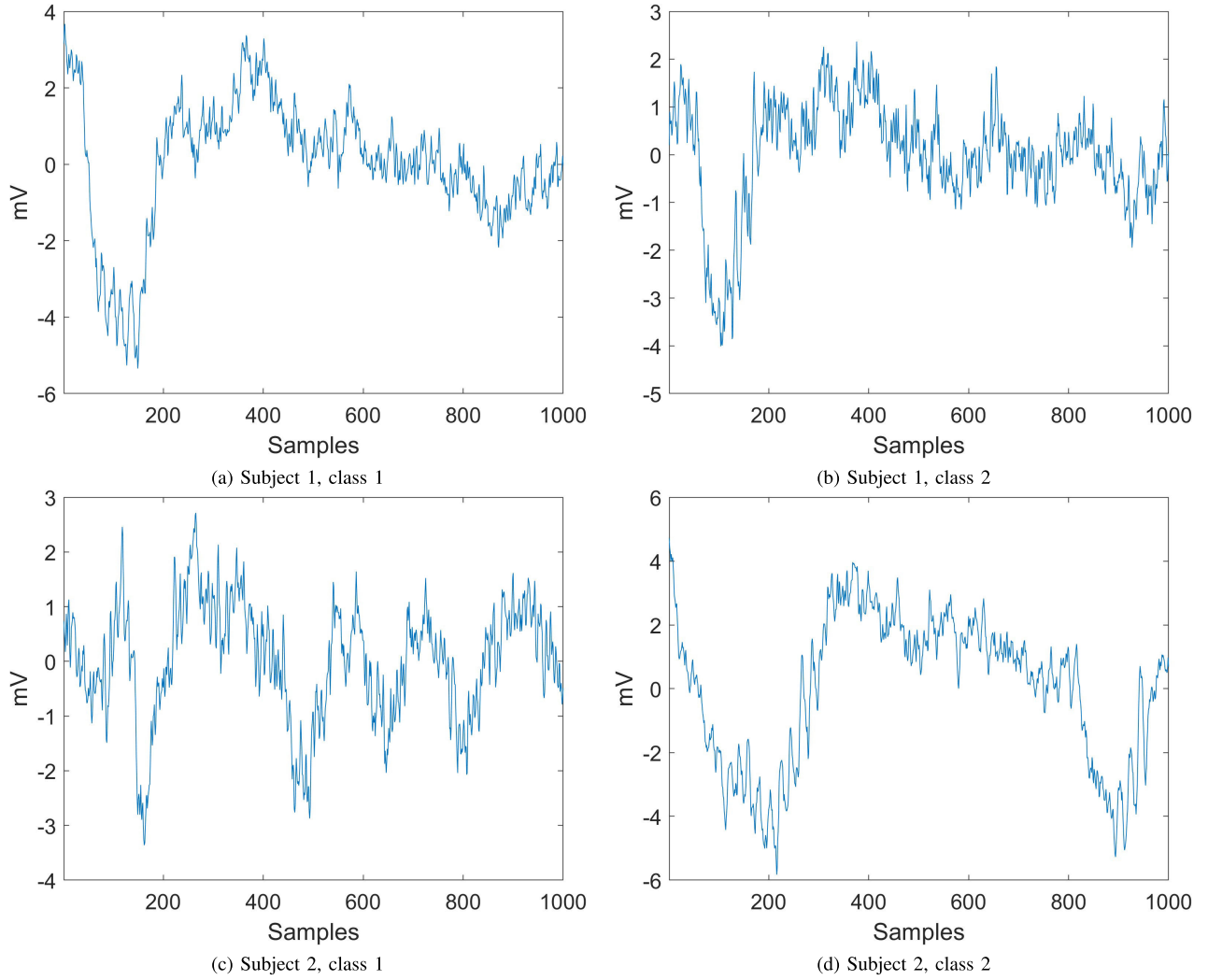


Fig. 1. EEG signals of two subjects performing motor imagery trials, recorded at channel 12, where classes 1 and 2 indicate the grasping of left and right hands, respectively.

clusters $c > 2 \in \mathbb{Z}$, and a fuzzy exponent $\beta > 1 \in \mathbb{R}$, the FCM attempts to optimally divide \mathbf{Y} into c partitions that are represented with c cluster centers by minimizing the following objective function:

$$F_\beta = \sum_{k=1}^N \sum_{l=1}^c (\mu_{kl})^\beta \|\mathbf{y}_k - \mathbf{v}_l\|^2, \quad (7)$$

where μ_{kl} is a short notation for $\mu(\mathbf{y}_k, \mathbf{v}_l)$, and the above objective function is subject to

$$\sum_{l=1}^c \mu_{kl} = 1, k = 1, \dots, N. \quad (8)$$

A numerical solution for minimizing F_β is by an iterative procedure that uses an initialization of the fuzzy membership to calculate the initial set of cluster centers. Mathematically, using the initialized fuzzy membership grades, the cluster centers are estimated as [33]

$$\mathbf{v}_l = \frac{\sum_{k=1}^N (\mu_{kl})^\beta \mathbf{y}_k}{\sum_{k=1}^N (\mu_{kl})^\beta}, \forall l. \quad (9)$$

The previous fuzzy membership grades are updated by

$$\mu_{kl} = \frac{1}{\sum_{d=1}^c \left(\frac{\|\mathbf{y}_k - \mathbf{v}_l\|}{\|\mathbf{y}_k - \mathbf{v}_d\|} \right)^{2/(\beta-1)}}, \forall k \forall l. \quad (10)$$

The process repeatedly performs the updating using Equations (9) and (10) until convergence or a defined maximum number of iterations are reached.

B. Wavelet Scattering Decompositions

The wavelet scattering decomposes the original data through a number of stages or layers of a tree structure, in which the output from one layer becomes the input for the next layer. Each layer involves three basic operations of transform: convolution (wavelet), nonlinearity (modulus), and average pooling (scaling function or low-pass filtering). The decomposition process for computing wavelet scattering coefficients is described as follows.

Let $\psi(t)$ be a band-pass filter with a center frequency normalized to 1 (mother wavelet), and $\psi_{\lambda_j}(t)$ be a wavelet filter bank constructed by dilating the mother

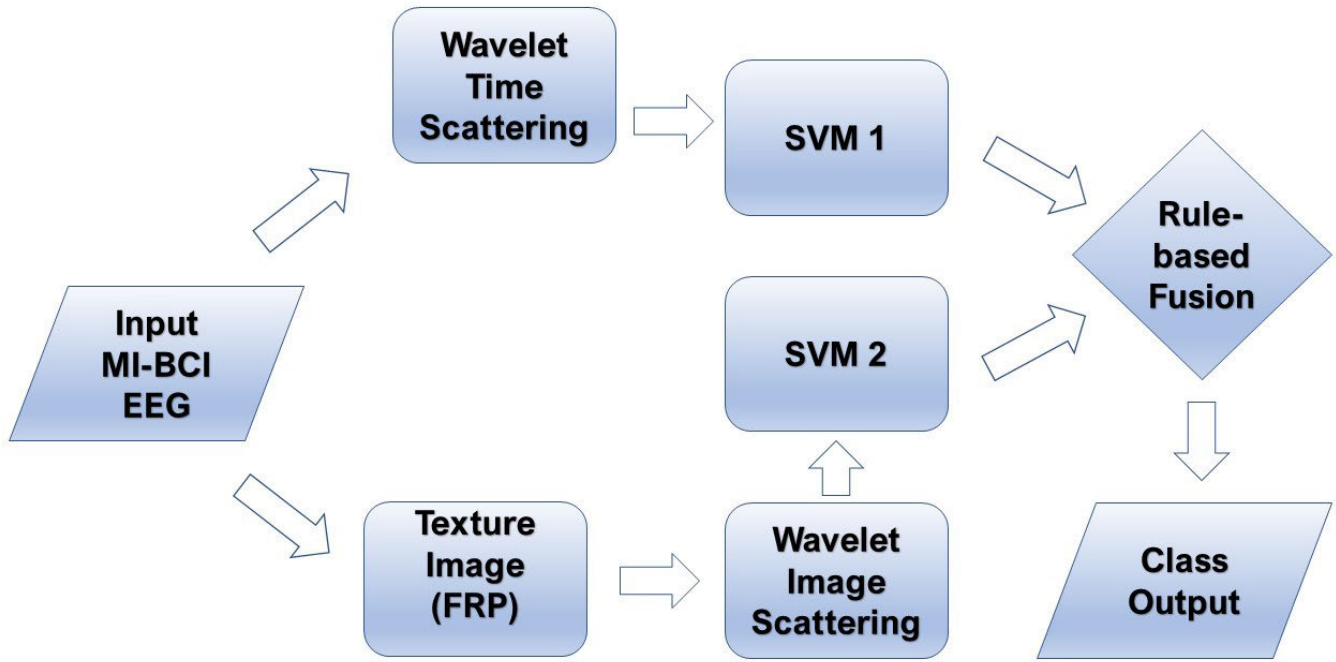


Fig. 2. Pipeline procedure for fusion of complimentary weak classifiers (SVMs) using wavelet time and image scattering features of MI-BCI EEG data.

wavelet as

$$\psi_{\lambda_j}(t) = \lambda_j \psi(\lambda_j t), \quad (11)$$

where $\lambda_j = 2^{(j/Q)}$, $j \in \mathbb{Z}$, $1 \leq j \leq J$, J is the maximum level of layers, and Q is the number of wavelets per octave.

Given an input signal \mathbf{u} , the wavelet scattering coefficients at layer zero (zeroth-order scattering coefficients), denoted as S_0 , are computed by the averaging of the signal as

$$S_0 = \mathbf{u} * \phi(t), \quad (12)$$

where ϕ is a scaling function or low-pass filter, and $*$ denotes the convolution operator.

The wavelet scattering at layer 1 or the 1st-order wavelet scattering coefficients are computed by averaging the modulus of the wavelet coefficients using ϕ as

$$S_1 = |\mathbf{u} * \psi_{\lambda_1}(t)| * \phi(t). \quad (13)$$

The 2nd-order wavelet scattering coefficients are calculated in a repeated manner in the sequential terms of convolution, modulus, and average pooling as

$$S_2 = ||\mathbf{u} * \psi_{\lambda_1}(t)| * \psi_{\lambda_2}(t)| * \phi(t). \quad (14)$$

In general, the wavelet scattering coefficients at layer j are determined as

$$S_j = |\dots ||\mathbf{u} * \psi_{\lambda_1}(t)| * \psi_{\lambda_2}(t)| \dots * \psi_{\lambda_j}(t)| * \phi(t). \quad (15)$$

In this study, the Morlet wavelet, which is the re-expression of the Gabor wavelet and widely used for determining wavelet scattering coefficients, is adopted as the mother wavelet ψ . The Morlet wavelet is defined as [34] and [35]

$$\psi(t) = c \left[e^{-\frac{t^2}{2\sigma^2}} \right] e^{2\pi i f t}, \quad (16)$$

where, in this study, c is a constant taken as 1, σ is the wavelet duration taken as 1, i is the imaginary unit, f is the center frequency, and $2\pi f = 5$.

For computing wavelet scattering coefficients of image data, a 2-D directional wavelet can be obtained by rotating a band-pass filter ψ by a number of rotations K having angles $2k\pi/K$ and dilating it by 2^j , such as [25]

$$\psi_{2^{-j}r}(t) = 2^{-2j} \psi(2^{-j} r^{-1} t), \quad (17)$$

where $r \in \{2q\pi/K\}$, $0 < q < K$.

C. Fusion of Complimentary Weak Classifiers

Suppose for a two-class prediction problem that two classifiers, which are support vector machines (SVMs) in this study, denoted as \mathcal{F}_1 and \mathcal{F}_2 , yield low classification accuracies from validation data. \mathcal{F}_1 well correctly predicts a certain class, denoted as C_1 , while highly misclassifies the other class, denoted as C_2 . \mathcal{F}_2 predicts the other way around, which is high for C_2 and low for C_1 . Such different predictions are complementary and can be used for weak classifier fusion. Based on the scores or probabilities computed for each class by the classifiers, denoted as s , a fusion strategy is designed using decision rules for classifying an input z as follows:

- Rule 1: If $\mathcal{F}_1(z) \mapsto C_1$ and $\mathcal{F}_2(z) \mapsto C_1$: Assign z to C_1 .
- Rule 2: If $\mathcal{F}_1(z) \mapsto C_2$ and $\mathcal{F}_2(z) \mapsto C_2$: Assign z to C_2 .
- Rule 3: If $\mathcal{F}_1(z) \mapsto C_1$ and $\mathcal{F}_2(z) \mapsto C_2$: Assign z to C_1 if $s[\mathcal{F}_1(z) \mapsto C_1] \geq s[\mathcal{F}_2(z) \mapsto C_2]$; otherwise assign z to C_2 .
- Rule 4: If $\mathcal{F}_1(z) \mapsto C_2$ and $\mathcal{F}_2(z) \mapsto C_1$: Assign z to C_2 if $s[\mathcal{F}_1(z) \mapsto C_2] \geq s[\mathcal{F}_2(z) \mapsto C_1]$; otherwise assign z to C_1 .

While the decisions in Rules 1 and 2 are based on consensus on the outputs of two classifiers, the decisions in Rules 3 and 4 aim to reduce biases of the two complimentary

weak classifiers toward two classes, respectively. In other words, Rule 3 attempts to reduce the misclassification of C_2 samples predicted as C_1 by \mathcal{F}_1 ; whereas Rule 4 tries to rectify the amount of C_1 samples misclassified as C_2 by \mathcal{F}_2 .

Figure 2 shows the graphical procedure for the classification of MI actions in BCI by the rule-based fusion of two complementary weak SVMs that are trained with wavelet time and image scattering features of the EEG signals.

IV. EXPERIMENTS

A. Parameter Settings

The EEG signals of 1000 samples in length for variable number of trials around 100, where each was recorded with 32 channels, of 5 sessions obtained from 25 subjects were used to extract the decomposition coefficients of wavelet time scattering by constructing a scattering network with two filter banks. The first filter bank has a quality factor $Q = 8$ wavelets per octave, and the second filter bank has a quality factor $Q = 1$ wavelet per octave. Other properties for creating the network of wavelet time scattering are: signal length = 1000, scattering transform invariance scale = 2 seconds, signal extension method = periodic, and sampling frequency = 250 Hz. Figure 3 shows the first and second filter bank wavelet filters for the wavelet time scattering constructed using the above specified properties.

To compute FRPs for the EEG data, the signal of only the first 202 samples of only 1 channel recorded by sensor number 12 located at the center of the scalp was used for constructing the FRP for the 5 sessions performed by the 25 subjects. Using an embedding dimension $m = 3$, time delay $\tau = 1$, the number of clusters $c = 3$, and fuzzy exponent $\beta = 2$, the transformation resulted in grayscale images of 200×200 pixels. Currently there are no analytical methods for finding optimal criteria for constructing an FRP in terms of m , τ , c , and β . The selected values were aimed to differentiate EEG signals of the two classes at a higher dimension ($m = 3$) with the smallest time separation ($\tau = 1$) between the occurrence of two events. The false nearest neighbor (FNN) and average mutual information (AMI) were developed for estimating parameters m and τ , respectively [36]. However, the estimations of these two parameters using FNN and AMI for the phase-space reconstruction of each EEG signal are deemed not effective for feature extraction because the signals may result in features with different dimensions. This issue still remains open for research in recurrence analysis [37], and it has been reported that optimal selections of values for m are problem-dependent [38]. Having discussed in [39], the FCM, which is adopted for constructing an FRP, is governed by two input parameters c and β . The questions of choosing appropriate values for c with problems underlying by an undefined number of clusters and β are still an on-going research area of fuzzy cluster analysis, where β was suggested to take real values between 1.5 and 2.5 and $\beta = 2$ has been the most widely adopted value in various applications of the FCM [40]. It was illustrated that different values of c appear to be insensitive in computing FRPs [30]. Previous works using the same values for m and β specified in this study for

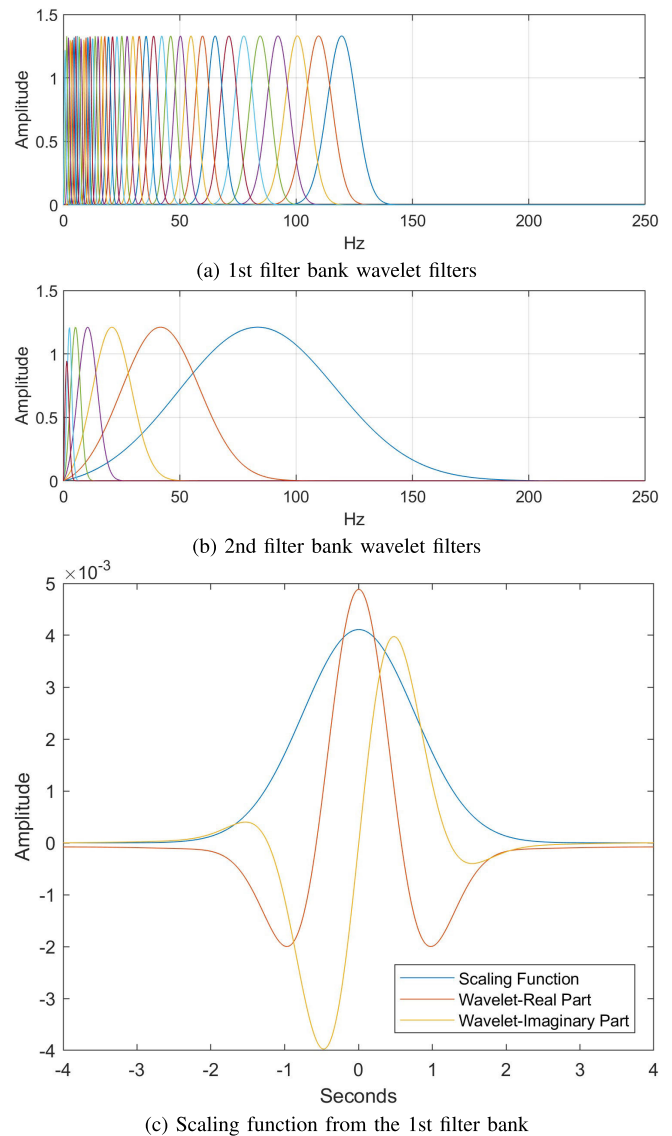


Fig. 3. Wavelet time scattering filters and the invariance scale.

determining the FRPs of physiological signals were found to be appropriate for the task of pattern classification [41], [42]. Figure 4 shows examples of EEG segments and corresponding texture images of the FRPs of two subjects.

To extract the wavelet scattering coefficients from the FRP images, the following parameters were used: image size = 200×200 , scattering transform invariance scale = 150 seconds, number of rotations per wavelet for two filter banks = 6, scattering filter bank quality factors (number of wavelet filters per octave) = 1, and a scattering path is carried out only if the bandwidth of the parent node overlaps significantly with that of the child node. Figure 5 shows the wavelet image scattering function from the first filter bank, and invariance scale constructed using the above specified parameters.

Parameters for the SVM-based classifier using the wavelet time scattering coefficients computed directly from the EEG signals are: kernel function = quadratic polynomial, using standardized data, and coding design = one versus one. For the SVM-based classifier using the wavelet image scattering

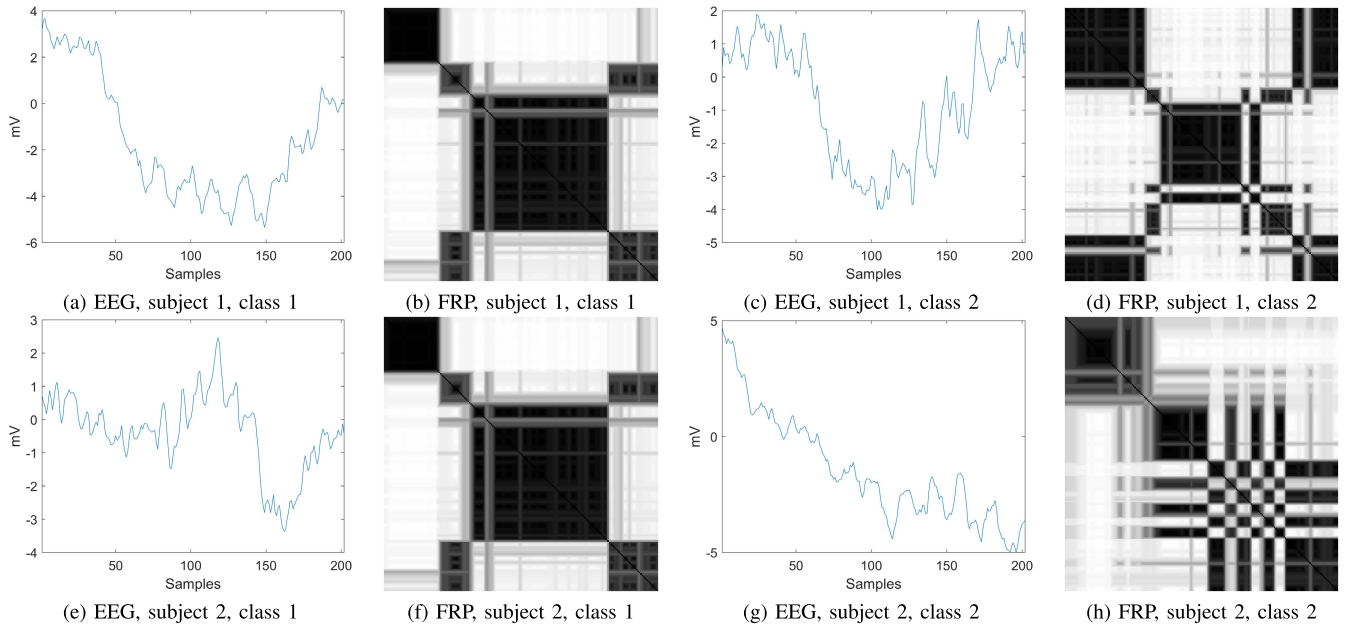


Fig. 4. Segments of first 202 samples of MI-BCI EEG signals and corresponding 200×200 texture images of FRPs.

TABLE I
MEASURES OF CLASSIFICATION PERFORMANCE

ACC	SEN	SPE	PRE	F_1
$\frac{TP + TN}{P + N}$	$\frac{TP}{P}$	$\frac{TN}{N}$	$\frac{TP}{TP + FP}$	$\frac{2TP}{2TP + FP + FN}$

coefficients computed from the FRPs: kernel function = linear, using unstandardized data, and coding design = one versus one.

To compare the proposed approach with a deep-learning model other than those reported in [5], both long short-term memory (LSTM) and bidirectional long short-term memory (Bi-LSTM) networks [43], [44] were used to learn on and classify the same training and test data, respectively. LSTM networks are a popular recurrent deep-learning approach for classifying sequential data or time series because these networks can learn long-term dependencies between time steps of the data. A Bi-LSTM is a recurrent neural network that learns order dependence in sequential data in both directions (backward and forward). LSTM has been used in BCI for decoding gait phases from EEG signals during locomotion [45]. Parameter specifications for LSTM and Bi-LSTM were: network layer with an output size = 100, fully connected layer = 2 (two classes), followed by a softmax layer and a classification layer. For training options, optimizer = ‘Adam’ (adaptive moment estimation), including L_2 regularization factor, maximum number of epochs = 40, minimum batch size = 150, initial learning rate = 0.01, and gradient threshold = 1.

B. Majority Vote

For the classification using the wavelet time scattering features, because for each EEG signal in this study, there

are 8 separate time windows (number of samples obtained after downsampling) and 32 channels for classification, the majority vote was applied for the two cases to obtain a single class prediction for each scattering representation and the set of channels, where there were equal numbers of votes for classifying the signal, a tie was recorded. For the classification using the wavelet image scattering features, because of using only a single channel, the majority vote was applied to obtain the classification for each scattering representation.

C. Measures of Performance

To obtain measures of classification results, accuracy (ACC), specificity (SPE), sensitivity (SEN), precision (PRE), and F_1 score are used in this study. Mathematical expressions for these statistical measures are given in Table I, which are computed in terms of condition positive (P), condition negative (N), true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In this study, P and N indicate the total numbers of class-2 (right-hand grasping) and class-1 (left-hand grasping) samples, respectively, in the data. TP and TN are the numbers of class-2 and class-1 samples, respectively, correctly predicted by a classifier. FP and FN are the numbers of class-2 and class-1 samples misclassified as class 1 and class 2, respectively, by a classifier. Thus, SPE and SEN are interpreted as true rates of class-1 and class-2 samples being correctly predicted by a classifier, ACC is the rate of both class-1 and class-2 samples being correctly classified. PRE is the proportion of class-2 samples that are truly predicted, and F_1 score is the harmonic mean of PRE and SEN.

D. Within-Session Classification Results

For the within-session classification, EEG signals obtained from all trials performed by the 25 participants were used for

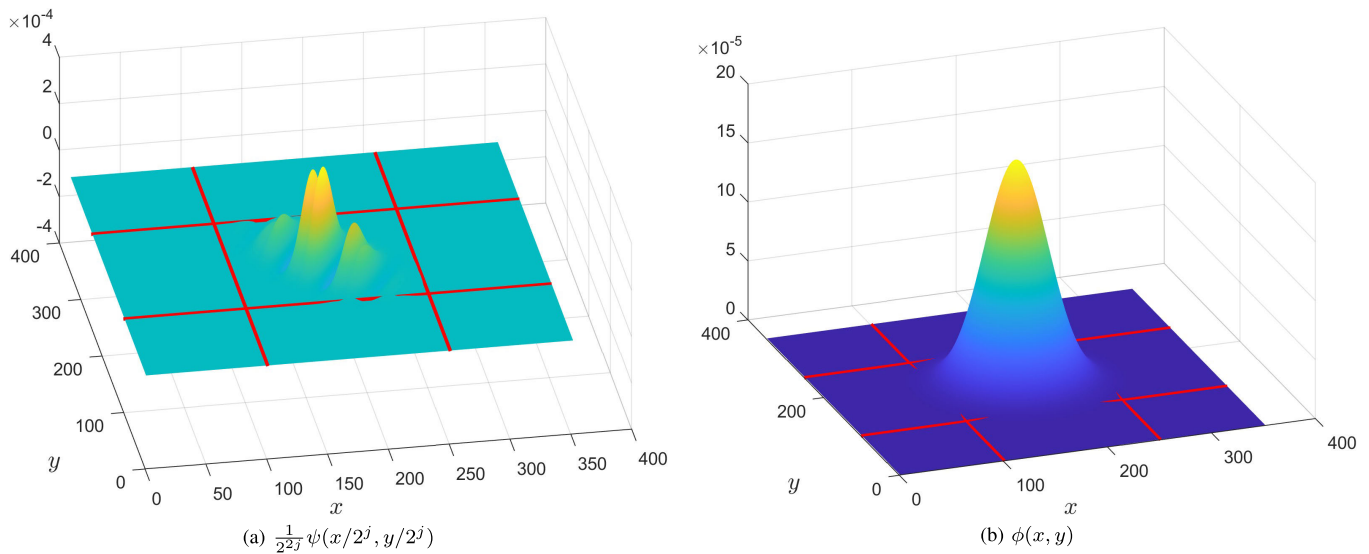


Fig. 5. Wavelet image scattering function from the first filter bank (a), and image invariance scale (b).

TABLE II
TEN-FOLD CROSS-VALIDATION RESULTS OF WITHIN-SESSION CLASSIFICATION OBTAINED FROM DIFFERENT CLASSIFIERS, WHERE NP = NOT PROVIDED, N = NO, AND Y = YES

Model	ACC (%)	SPE (%)	SEN (%)	PRE (%)	F_1	$p < 0.01$
CSP [5]	57.33	NP	NP	NP	NP	N
FBCSP [5]	64.33	NP	NP	NP	NP	N
FBCNet [5]	68.85	NP	NP	NP	NP	N
EEGNet [5]	65.03	NP	NP	NP	NP	N
deep ConvNets [5]	64.82	NP	NP	NP	NP	N
LSTM	59.66	60.06	59.27	59.94	0.60	Y
Bi-LSTM	61.83	61.31	62.32	62.26	0.62	Y
WIS-SVM	50.66	74.26	25.60	25.60	0.23	Y
WTS-SVM	65.51	54.52	76.94	76.94	0.67	Y
WTIS-SVM	75.68	74.26	76.94	76.94	0.75	Y

carrying out a 10-fold cross-validation (CV), which randomly split the data into 10 equal parts, used 9 of those parts for extracting wavelet-time scattering features and wavelet-image scattering features for training the WTS-SVM and WIS-SVM, respectively, and the remaining part for testing. The validation was repeated 10 times, where a different tenth part was selected for testing each time. The results were the average of the 10 runs of the 10-fold cross-validation.

Table II shows the results computed using a 10-fold CV obtained from different conventional and advanced AI classification models using various types of features: common spatial patterns (CSP) [46], filter bank common spatial pattern (FBCSP) [47], filter-bank convolutional network (FBCNet) [48], EEGNet [49], deep convolutional network (deep ConvNets) [50], LSTM, Bi-LSTM, wavelet image scattering-based support vector machine (WIS-SVM), wavelet time scattering-based support vector machine (WTS-SVM), and fusion of wavelet time and image scattering-based support vector machines (WTIS-SVM). Figure 6 shows the training progresses of the LSTM and Bi-LSTM, in which convergences were reached at early stages. The 10-fold CV results produced from CSP, FBCSP, FBCNet, EEGNet, and deep ConvNets as shown in Table II were originally reported in reference [5].

TABLE III
95% CONFIDENCE INTERVALS (CIs) FOR THE MEAN ACCURACIES COMPUTED FROM THE t -TEST OF WITHIN-SESSION CLASSIFICATION RESULTS GIVEN BY SVM-BASED CLASSIFIERS USING WAVELET SCATTERING FEATURES OF TIME AND IMAGE DATA, AND FUSION OF CLASSIFIERS

Model	CI	p -value
WIS-SVM	[49.98, 51.34]	1.48e-93
WTS-SVM	[63.35, 67.67]	3.23e-141
WTIS-SVM	[71.75, 79.62]	3.16e-70

Table III shows the 95% confidence intervals of the mean accuracies computed from the t -test of 10-fold CV obtained from the SVM-based classifiers using wavelet scattering features of the EEG signals, FRPs of the EEG signals, and fusion of two SVM-based classifiers.

Both WIS-SVM and WTS-SVM are weak classifiers, whose average accuracies are 61% and 66%, respectively. Correct predictions of class 1 (SPE) and class 2 (SEN) provided by both models are opposite. SPE obtained from the WIS-SVM (74%) is much higher than the WTS-SVM, but SEN obtained from the WTS-SVM (77%) is much higher than the WIS-SVM. Such prediction outcomes are complementary

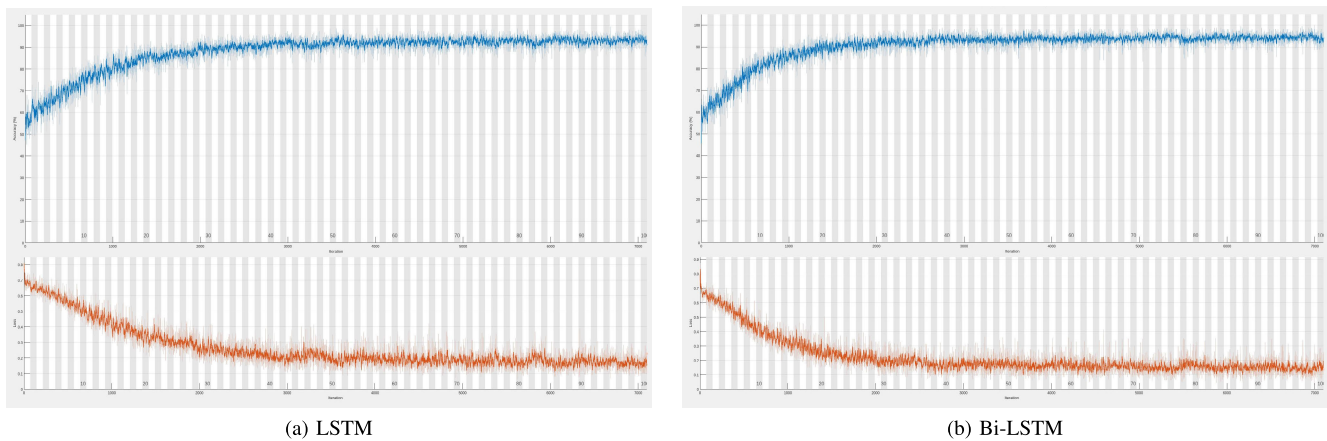


Fig. 6. Training progresses of LSTM and Bi-LSTM networks for within-session 10-fold cross-validation, where upper and lower y-axes are accuracy and loss, respectively, and x-axes are the number of iterations.

for classifier fusion. The Bi-LSTM (ACC = 62%) performed slightly better than the LSTM (ACC = 60%). As a result, the combined model WTIS-SVM yielded the highest classification accuracy (76%) over the other 7 classifiers as shown in Table II. SPE, SEN, PRE, and F_1 values provided by the WITS-SVM are balanced, which are 74% and 77%, 77%, and 0.75, respectively. The 10-fold CV accuracies obtained from other methods (CSP, FBCSP, FBCNet, EEGNet, and deep ConvNets) reported in a previous study [5] as shown in Table II do not support the 1% statistical significance level ($p \not\leq 0.01$), whereas those obtained from WIS-SVM, WTS-SVM, and WTIS-SVM do.

E. Cross-Session Classification Results

For the cross-session classification, all sessions of each participant were selected as the test data or target domain, and all sessions of the other 24 participants were used for training or treated as the source domain. This type of classification aims to adapt trained models to recognize MI-based tasks of new subjects, which are more useful for BCI applications.

Table IV shows the classification results obtained from the EEGNet [5], deep ConvNets [5], and FBCNet [5], LSTM, Bi-LSTM, WIS-SVM, WTS-SVM, and WTIS-SVM. Table III shows the 95% confidence intervals of the mean accuracies computed using the t -test for the WIS-SVM, WTS-SVM, and WTIS-SVM models. Figure 7 shows the training progresses of the LSTM and Bi-LSTM, in which convergences were reached at early stages. Both Figure 6 and Figure 7 show a strong similarity in the learning the EEG signals.

For individual classifiers, the WTS-SVM achieved the best accuracy (ACC = 59%). The Bi-LSTM model yielded the second-best performance (ACC = 53.91%), which is slightly better than the EEGNet (53.65%). EEGNet, deep ConvNet, and Bi-LSTM are competitive in terms of accuracy. Overall, classification accuracies obtained all four individual classifiers (EEGNet, deep ConvNets, FBCNet, LSTM, Bi-LSTM, WIS-SVM, and WTS-SVM) were below 60%, reflecting the cross-session classification is a more challenging task than the within-session classification. Being similar to the case of within-session classification, the SPE obtained from the

WIS-SVM was higher than the WTS-SVM, and the SEN obtained from the WIS-SVM was lower than the WTS-SVM, showing a potential complement for classifier fusion. As a result, the proposed fusion model WTIS-SVM provided the highest ACC (65%) among the LSTM, Bi-LSTM, WIS-SVM, and WTS-SVM models. SPE and SEN values obtained from the WTIS-SVM are balanced (65% and 64%, respectively). The fusion model provided the best PRE (65%) and second highest F_1 (0.65) that is lower than F_1 value obtained from WTS-SVM (0.69).

Table V shows the 95% confidence intervals of the mean accuracies computed from the t -test for the cross-session classification obtained from the WTS-SVM, WIS-SVM, and WTIS-SVM classifiers.

F. Discussion

The classification based on the wavelet scattering decomposition of the FRPs was much faster than that of the EEG signals, because the former extracted the wavelet decomposition features from a much shorter signal (202 vs 1000 samples) from only 1 out of 32 channels. In this case, time-based and image-based wavelet scattering features of the EEG signals provided biased information about a particular class for the prediction task performed by the SVMs. By utilizing complementary results output from the two SVM-based models, the prediction of motor imagery from EEG signals could be improved using a simple fusion strategy. In fact, research into combining strengths of multiple weak classifiers for achieving more accurate results than those obtained from individual biased classifier has been explored and applied to many fields [51], [52], [53], [54], [55].

Sensitivity analysis among signal lengths and classification accuracy was also investigated for computing the wavelet scattering coefficients. Because the applied sampling frequency of 250 Hz and recorded EEG signal length of 1000 samples were fixed, there was no sensitivity analysis for computing the wavelet-time scattering coefficients. For the wavelet-image scattering coefficients determined by means of the FRPs of the original brain data, the length of the EEG signals was varied between 182 and 222 in the increment of 10 samples, yielding

TABLE IV
CROSS-SESSION CLASSIFICATION OBTAINED FROM DIFFERENT CLASSIFIERS, WITH $p < 0.01$

Model	ACC (%)	SPE (%)	SEN (%)	PRE (%)	F_1
EEGNet [5]	53.65	NP	NP	NP	NP
deep ConvNets [5]	52.92	NP	NP	NP	NP
FBCNet [5]	50.97	NP	NP	NP	NP
LSTM	49.46	42.70	56.33	49.89	0.46
Bi-LSTM	53.91	51.46	56.42	54.24	0.53
WIS-SVM	49.27	69.00	50.10	50.13	0.30
WTS-SVM	59.38	29.61	88.32	56.43	0.69
WTIS-SVM	64.84	65.23	64.43	65.44	0.65

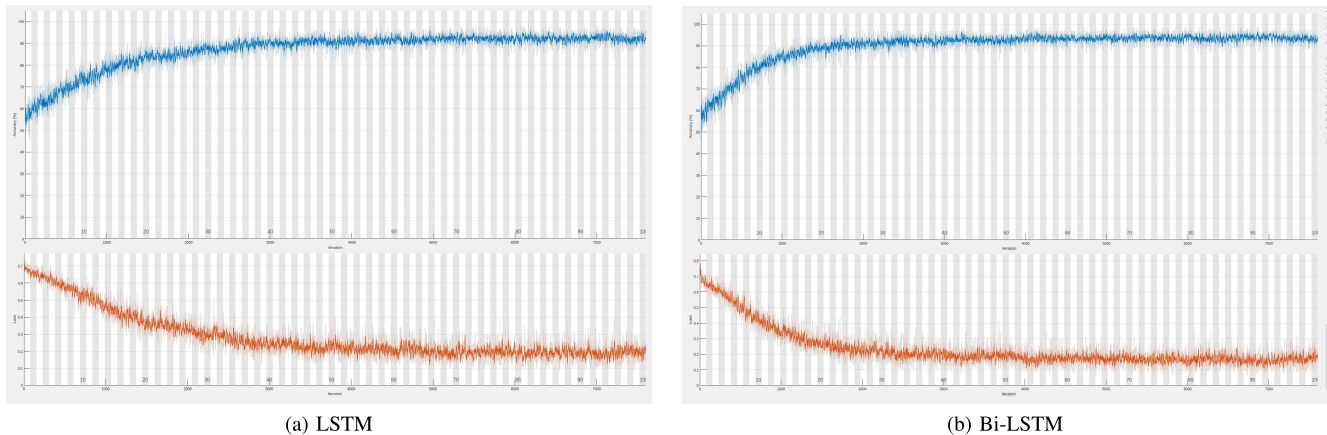


Fig. 7. Training progresses of LSTM and Bi-LSTM networks for cross-session classification, where upper and lower y-axes are accuracy and loss, respectively, and x-axes are the number of iterations.

TABLE V
95% CONFIDENCE INTERVALS (CIs) FOR THE MEAN ACCURACIES COMPUTED FROM THE t -TEST OF CROSS-SESSION CLASSIFICATION RESULTS GIVEN BY SVM-BASED CLASSIFIERS USING WAVELET SCATTERING FEATURES OF TIME AND IMAGE DATA, AND FUSION OF CLASSIFIERS

Model	CI	p -value
WIS-SVM	[46.08, 52.47]	1.78e-06
WTS-SVM	[52.86, 65.91]	1.46e-05
WTIS-SVM	[60.01, 69.66]	3.08e-06

FRPs of 180×180 , 190×190 , 200×200 , 210×210 , and 220×200 grayscale images. The average accuracies obtained from WIS-SVM and WTIS-SVM for the within-session classification were 51.79% and 75.77%, respectively, which are only slightly higher than those listed in Table II (50.66% and 75.68%, respectively). The average accuracies obtained from WIS-SVM and WTIS-SVM for the cross-session classification were 48.92% and 64.69%, respectively, which are very slightly lower than those listed in Table IV (49.27% and 64.84%, respectively). No clear linear relationships between signal lengths and accuracies were found in both within-session and cross-session tests.

Advantages of the method introduced in this study are that (1) the fusion strategy is general, which is not limited to the use of SVMs but can be applied to any types of classifiers whose predictive outputs are complementary by checking the confusion matrices of training data; (2) the use of short segments of the physiological signals for transforming

1-D into 2-D texture data for extracting features from wavelet image scattering decomposition allows efficient computation in both speed and memory storage, which is very useful for processing big data; and (3) the transformation of shorts segments of the EEG signals recorded from only one out of 32 sensors into FRPs that produced complementary results appears to offer an economic solution in terms of device setup and convenience to participants undertaking rehabilitation training.

Disadvantages of the proposed approach include the increase in computational complexity conditioned by the transformation of EEG signals into grayscale images of FRPs, and the need for optimal parameter setting for constructing the FRPs, which has not been analytically found in the present study. Furthermore, the fuzzy rule-based system can be effective only if the sensitivity and specificity obtained from the WIS-SVM and WTS-SVM are complementary.

Being similar to the basic concept of the power spectrum, which is the mathematical expression of the quantity of the signal at a certain frequency, the idea of wavelet analysis is to compute how much a wavelet is contained in a signal for a certain scale and location. Particularly for texture analysis, wavelet image scattering constructs low-variance representations of the FRPs of the EEG signals, which are insensitive to translations. The process of cascading the textural FRPs through a series of transforms, nonlinearities, and averaging in the wavelet scattering was found to provide biased features toward class 1 of the MI-based EEG signals. On another aspect, wavelet time scattering resulted in the transformation of the EEG signals being insensitive to shifts in the physiologic

signals, which was observed to favor the prediction of class 2 of the visual BCI experiments. Results obtained from the two SVM-based classifiers trained with two different types of wavelet scattering features suggest different characteristic patterns of the signals can be captured in time and space domains. These characteristics can be used to enhance the task of classification by means of a strategy for selecting the outputs from the classifiers. The CV measures of performance produced by combining the two weak SVM-based classifiers are higher than those provided by other classic MI-BCI algorithms and deep-learning models, which are among state-of-the-art AI methods.

The fusion of two SVM-based classifiers, where one learned on wavelet time scattering and the other on wavelet image scattering of the EEG signals, have been shown to perform better than several deep-learning models. A main reason for the higher performance of the fusion model is that the two SVM-based classifiers are complementary to each other by learning on temporal and spatial features of the wavelet transform, and the use of the fuzzy rules was able to rectify a number of misclassified signals by targeting on different strengths of the two classifiers. To elaborate further on the wavelet scattering, the wavelet time scattering created a network for wavelet time scattering decomposition of the EEG data using the analytic Morlet wavelet. The network then used wavelets and a low-pass filtering function to generate low-variance representations of the physiological signals, yielding extracted features robust to time translations in the input signals. On the other hand, wavelet image scattering constructed low-variance representations of the fuzzy recurrence plots of the EEG signals, which are insensitive to spatial translations. Thus, machine learning on these two types of wavelet-scattering features of the EEG signals are expected to be complementary. There are reports showing that SVMs performed competitively or better than deep learning in several cases of pattern recognition [56], [57], [58], [59].

V. CONCLUSION

The proposed complimentary weak classifier fusion for classifying MI tasks in BCI using a big challenging database of EEG signals has been presented and discussed in the foregoing sections. The novel aspects of the proposed approach include: 1) the exploration of wavelet scattering networks of signals and images for extracting complementary features of EEG data in BCI to be learned with the SVM has been presented herein for the first time; 2) the FRP construction allows the EEG signals to be transformed into grayscale images, discovering a new dimension of features of the original brain signals; 3) the development of a fuzzy rule-based system has been illustrated effective for combining outputs from two complementary SVM-based classifiers; and 4) the experimental results obtained from the proposed fusion are shown to be better than several existing state-of-the-art classification methods. Furthermore, not only the within-session classification but also the inclusion of the cross-session classification of the MI activities, has been addressed in this study, showing a potential contribution to the practical aspects of MI-BCI using EEG data [5], [60].

In this study, the transformation of the EEG signals into texture images by means of the FRP algorithm appears to be capable of capturing differential information from short segments of the two different brainwave activities from a single electrode, and therefore worth further investigating in conjunction with other texture feature extraction methods for pattern classification. The selection of the channel location in this study was based on its central position on the head. Sensitivity analysis of different individuals and alternative combinations of the channels would be of interest for improving the classification. Training of deep-learning methods such as pretrained CNN models with the rich texture of FRPs, and LSTM networks with multiple wavelet time and image scattering features of the segmented or whole EEG signals would also be particular areas of interest. The combination of AI, signal processing, and data science is promising for advancing computerized control and communication technology for aiding the rehabilitation of patients with stroke or neurological disabilities.

REFERENCES

- [1] T. Mulder, "Motor imagery and action observation: Cognitive tools for rehabilitation," *J. Neural Transmiss.*, vol. 114, no. 10, pp. 1265–1278, Oct. 2007.
- [2] M. Lotze and L. G. Cohen, "Volition and imagery in neurorehabilitation," *Cogn. Behav. Neurol.*, vol. 19, no. 3, pp. 135–140, Sep. 2006.
- [3] S. Park, J. Ha, D.-H. Kim, and L. Kim, "Improving motor imagery-based brain-computer interface performance based on sensory stimulation training: An approach focused on poorly performing users," *Frontiers Neurosci.*, vol. 15, Nov. 2021, Art. no. 732545.
- [4] P. D. E. Baniqued et al., "Brain-computer interface robotics for hand rehabilitation after stroke: A systematic review," *J. NeuroEng. Rehabil.*, vol. 18, p. 15, Jan. 2021.
- [5] J. Ma, B. Yang, W. Qiu, Y. Li, S. Gao, and X. Xia, "A large EEG dataset for studying cross-session variability in motor imagery brain-computer interface," *Sci. Data*, vol. 9, no. 1, p. 531, Sep. 2022.
- [6] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophys.*, vol. 113, no. 6, pp. 767–791, 2002.
- [7] C. Neuper, G. R. Müller, A. Kübler, N. Birbaumer, and G. Pfurtscheller, "Clinical application of an EEG-based brain-computer interface: A case study in a patient with severe motor impairment," *Clin. Neurophysiol.*, vol. 114, no. 3, pp. 399–409, Mar. 2003.
- [8] J. S. Brumberg, A. Nieto-Castanon, P. R. Kennedy, and F. H. Guenther, "Brain-computer interfaces for speech communication," *Speech Commun.*, vol. 52, no. 4, pp. 367–379, 2010.
- [9] J. J. Daly and J. R. Wolpaw, "Brain-computer interfaces in neurological rehabilitation," *Lancet Neurol.*, vol. 7, no. 11, pp. 1032–1043, 2008.
- [10] R. Mane, T. Chouhan, and C. Guan, "BCI for stroke rehabilitation: Motor and beyond," *J. Neural Eng.*, vol. 17, no. 4, Aug. 2020, Art. no. 041001.
- [11] M. Sebastián-Romagosa et al., "Brain computer interface treatment for motor rehabilitation of upper extremity of stroke patients—A feasibility study," *Frontiers Neurosci.*, vol. 14, Oct. 2020, Art. no. 591435.
- [12] R. Foong, K. A. Kai, Q. Chai, C. Guan, and K. Chua, "Assessment of the efficacy of EEG-based MI-BCI with visual feedback and EEG correlates of mental fatigue for upper-limb stroke rehabilitation," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 3, pp. 786–795, Mar. 2020.
- [13] M. Alimardani, S. Nishio, and H. Ishiguro, "Brain-computer interface and motor imagery training: The role of visual feedback and embodiment," in: *Evolving BCI Therapy—Engaging Brain State Dynamics*. London, U.K.: IntechOpen, 2018. [Online]. Available: <https://www.intechopen.com/chapters/62072>
- [14] A. Roc et al., "A review of user training methods in brain computer interfaces based on mental tasks," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021, Art. no. 011002.
- [15] P. R. Kennedy and K. D. Adams, "A decision tree for brain-computer interface devices," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 148–150, Jun. 2003.

- [16] S. Aggarwal and N. Chugh, "Signal processing techniques for motor imagery brain computer interface: A review," *Array*, vols. 1–2, Jan. 2019, Art. no. 100003.
- [17] K. Zhang, N. Robinson, S. W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network," *Neural Netw.*, vol. 136, pp. 1–10, Apr. 2021.
- [18] A. M. Roy, "An efficient multi-scale CNN model with intrinsic feature integration for motor imagery EEG subject classification in brain-machine interfaces," *Biomed. Signal Process. Control*, vol. 74, Apr. 2022, Art. no. 103496.
- [19] N. Tibrewal, N. Leeuwis, and M. Alimardani, "Classification of motor imagery EEG using deep learning increases performance in inefficient BCI users," *PLoS ONE*, vol. 17, no. 7, Jul. 2022, Art. no. e0268880.
- [20] H. Zhu, D. Forenzo, and B. He, "On the deep learning models for EEG-based brain-computer interface using motor imagery," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2283–2291, 2022.
- [21] E. Dagdevir and M. Tokmakci, "Truncation thresholds based empirical mode decomposition approach for classification performance of motor imagery BCI systems," *Chaos, Solitons Fractals*, vol. 152, Nov. 2021, Art. no. 111450.
- [22] F. Wang et al., "Improved brain-computer interface signal recognition algorithm based on few-channel motor imagery," *Frontiers Hum. Neurosci.*, vol. 16, May 2022, Art. no. 880304.
- [23] S. Mallat, "Group invariant scattering," *Commun. Pure Appl. Math.*, vol. 65, no. 10, pp. 1331–1398, 2012.
- [24] J. Andén and S. Mallat, "Deep scattering spectrum," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4114–4128, Aug. 2014.
- [25] J. Bruna and S. Mallat, "Invariant scattering convolution networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1872–1886, Aug. 2013.
- [26] A. B. Buriro et al., "Classification of alcoholic EEG signals using wavelet scattering transform-based features," *Comput. Biol. Med.*, vol. 139, Dec. 2021, Art. no. 104969.
- [27] A. Sepúlveda, F. Castillo, C. Palma, and M. Rodríguez-Fernandez, "Emotion recognition from ECG signals using wavelet scattering and machine learning," *Appl. Sci.*, vol. 11, no. 11, p. 4945, May 2021.
- [28] Z. Liu, G. Yao, Q. Zhang, J. Zhang, and X. Zeng, "Wavelet scattering transform for ECG beat classification," *Comput. Math. Methods Med.*, vol. 2020, Oct. 2020, Art. no. 3215681.
- [29] L. Sifre and S. Mallat, "Rotation, scaling and deformation invariant scattering for texture discrimination," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1233–1240.
- [30] T. D. Pham, "Fuzzy recurrence plots," *Europhys. Lett.*, vol. 116, no. 5, p. 50008, Dec. 2016.
- [31] L. Ma, B. Yang, and W. Qiu, "SHU dataset," *Figshare*, Accessed: Sep. 8, 2022, doi: [10.6084/m9.figshare.19228725.v1](https://doi.org/10.6084/m9.figshare.19228725.v1).
- [32] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980* (Lecture Notes in Mathematics), vol. 898. Berlin, Germany: Springer-Verlag, 1981, pp. 366–381.
- [33] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. New York, NY, USA: Plenum, 1981.
- [34] J. Morlet, G. Arens, E. Fourgeau, and D. Glard, "Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media," *Geophysics*, vol. 47, no. 2, pp. 203–221, Feb. 1982.
- [35] J. Morlet, G. Arens, E. Fourgeau, and D. Giard, "Wave propagation and sampling theory—Part II: Sampling theory and complex waves," *Geophysics*, vol. 47, no. 2, pp. 222–236, Feb. 1982.
- [36] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [37] N. Marwan and K. H. Kraemer, "Trends in recurrence analysis of dynamical systems," *Eur. Phys. J. Special Topics*, pp. 1–23, Jan. 2023, doi: [10.1140/epjs/s11734-022-00739-8](https://doi.org/10.1140/epjs/s11734-022-00739-8).
- [38] K. H. Kraemer, M. Gelbrecht, I. Pavithran, R. I. Sujith, and N. Marwan, "Optimal state space reconstruction via Monte Carlo decision tree search," *Nonlinear Dyn.*, vol. 108, no. 2, pp. 1525–1545, Apr. 2022.
- [39] T. D. Pham, "From raw pixels to recurrence image for deep learning of benign and malignant mediastinal lymph nodes on computed tomography," *IEEE Access*, vol. 9, pp. 96267–96278, 2021.
- [40] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Trans. Fuzzy Syst.*, vol. 3, no. 3, pp. 370–379, Aug. 1995.
- [41] T. D. Pham, "Pattern analysis of computer keystroke time series in healthy control and early-stage Parkinson's disease subjects using fuzzy recurrence and scalable recurrence network features," *J. Neurosci. Methods*, vol. 307, pp. 194–202, Sep. 2018.
- [42] T. D. Pham, K. Wardell, A. Eklund, and G. Salerud, "Classification of short time series in early Parkinson's disease with deep learning of fuzzy recurrence plots," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 6, pp. 1306–1317, Nov. 2019.
- [43] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [44] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
- [45] S. Tortora, S. Ghidoni, C. Chisari, S. Micera, and F. Artoni, "Deep learning-based BCI for gait decoding from EEG with LSTM recurrent neural network," *J. Neural Eng.*, vol. 17, no. 4, Jul. 2020, Art. no. 046011.
- [46] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000.
- [47] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jun. 2008, pp. 2390–2397.
- [48] R. Mane, N. Robinson, A. P. Vinod, S.-W. Lee, and C. Guan, "A multi-view CNN with novel variance layer for motor imagery brain computer interface," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 2950–2953.
- [49] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.
- [50] R. T. Schirrmeister et al., "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Aug. 2017.
- [51] C. Ji and S. Ma, "Combinations of weak classifiers," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 32–42, Jan. 1997.
- [52] F. Parisi, F. Strino, B. Nadler, and Y. Kluger, "Ranking and combining multiple predictors without labeled data," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 4, pp. 1253–1258, 2014.
- [53] M. Komlagan et al., "Anatomically constrained weak classifier fusion for early detection of Alzheimer's disease," *Machine Learning in Medical Imaging* (Lecture Notes in Computer Science), vol. 8679. New York, NY, USA: Springer, 2014, pp. 141–148.
- [54] D. Miao, M. Zhang, H. Li, Z. Sun, and T. Tan, "Bin-based weak classifier fusion of iris and face biometrics," in *Proc. IEEE 7th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Sep. 2015, pp. 1–6.
- [55] Y. Hon and E. Yang, "Unbiased classification through bias-contrastive and bias-balanced learning," in *Proc. 35th Conf. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, 2021, pp. 26449–26461.
- [56] P. Liu, K.-K. R. Choo, L. Wang, and F. Huang, "SVM or deep learning? A comparative study on remote sensing image classification," *Soft Comput.*, vol. 21, no. 23, pp. 7053–7065, 2017.
- [57] J. Smolander, M. Dehmer, and F. Emmert-Streib, "Comparing deep belief networks with support vector machines for classifying gene expression data from complex disorders," *FEBS Open Bio*, vol. 9, no. 7, pp. 1232–1248, Jul. 2019.
- [58] P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognit. Lett.*, vol. 141, pp. 61–67, Jan. 2021.
- [59] H. Shirmard et al., "A comparative study of convolutional neural networks and conventional machine learning models for lithological mapping using remote sensing data," *Remote Sens.*, vol. 14, no. 4, p. 819, Feb. 2022.
- [60] V. Peterson et al., "Transfer learning based on optimal transport for motor imagery brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 2, pp. 807–817, Feb. 2022.