A Deep Learning Framework for Decoding Motor Imagery Tasks of the Same Hand Using EEG Signals

RAMI ALAZRAI[1], MOTAZ ABUHIJLEH[1], HISHAM ALWANNI[2], AND MOHAMMAD I. DAOUD[1]

[1]Department of Computer Engineering, School of Electrical Engineering and Information Technology, German Jordanian University, Amman 11180, Jordan
[2]Faculty of Engineering, University of Freiburg, 79098 Freiburg im Breisgau, Germany

Corresponding author: Rami Alazrai (rami.azrai@gju.edu.jo)

ABSTRACT This study aims to increase the control's dimensions of the electroencephalography (EEG)-based brain-computer interface (BCI) systems by distinguishing between the motor imagery (MI) tasks associated with fine body-parts of the same hand, such as the wrist and fingers. This in turn can enable individuals who are suffering from transradial amputations to better control prosthetic hands and to perform various dexterous hand tasks. In particular, we present a novel three-stage framework for decoding MI tasks of the same hand. The three stages of the proposed framework are the input, feature extraction, and classification stages. At the input stage, we employ a quadratic time-frequency distribution (QTFD) to analyze the EEG signals in the joint time-frequency domain. The use of a QTFD enables to transform the EEG signals into a set of two-dimensional (2D) time-frequency images (TFIs) that describe the distribution of the energy encapsulated within the EEG signals in terms of the time, frequency, and electrode position. At the feature extraction stage, we design a new convolutional neural network (CNN) architecture that can automatically analyze and extract salient features from the TFIs created at the input stage. Finally, the features obtained at the feature extraction stage are passed to the classification stage to assign each input TFI to one of the eleven MI tasks that are considered in the current study. The performance of our proposed framework is evaluated using EEG signals that were acquired from eighteen able-bodied subjects and four transradial amputated subjects while performing eleven MI tasks within the same hand. The average classification accuracies obtained for the able-bodied and transradial amputated subjects are 73.7% and 72.8%, respectively. Moreover, our proposed framework yields 14.5% and 11.2% improvements over the results obtained for the able-bodied and transradial amputated subjects, respectively, using conventional QTFD-based handcrafted features and a multi-class support vector machine classifier. The results demonstrate the efficacy of the proposed framework to decode the MI tasks associated with the same hand for able-bodied and transradial amputated subjects.

INDEX TERMS Convolutional neural networks (CNN), deep learning, electroencephalography (EEG), motor imagery, time-frequency distribution.

I. INTRODUCTION

Transradial amputations can profoundly reduce the quality of life of affected individuals [1]. This can be attributed to the fact that individuals who are suffering from transradial amputations are incapable of performing various dexterous hand movements that are necessary for many essential activities of daily living. Recently, many remarkable advances have been witnessed in developing dexterous upper limb robotic prosthetics, such as robotic prosthetic hands [2], [3]. These prostheses have the potential to enable individuals with transradial amputations to restore a significant part of their missing limbs [4]. In this regard, brain-computer interfaces (BCI) systems, which analyze brain activity and translate it into control commands, have been designed and employed to increase the control dimensions of the existing upper limb prostheses [5].

Among the different techniques used to record brain activities, the electroencephalography (EEG) technique has been

The associate editor coordinating the review of this article and approving it for publication was Lei Ding.

widely employed in designing BCI systems [5]–[7]. This can be attributed to the noninvasive nature of the EEG technique, high portability, low cost, and high temporal resolution of the recorded brain activities [5]. Different types of EEG signals have been utilized in EEG-based BCI systems, such as the steady-state visual evoked potentials, P300 evoked potentials, and motor imagery (MI). In this work, we focus on analyzing brain activities related to MI tasks in order to generate control commands for prosthetic hands. In particular, a MI task is the process of imagining to perform a specific task without actually performing it [8], [9]. Such a process involves the activation of different regions in the brain, including the primary motor cortex , dorsal and ventral pre-motor cortices, primary and secondary sensory areas, and pre-frontal areas [9], [10]. Therefore, developing EEG-based BCI systems that can accurately distinguish between different MI tasks is crucial to enhance the control of neural-based prosthetic hands, which in turn can improve the quality of life for individuals who are suffering from transradial amputations.

Literature reveals that the majority of the existing MI studies have focused on developing EEG-based BCI systems that can distinguish between the MI tasks related to five different large body-parts [6], [11]–[17], namely the left hand, right hand, left foot, right foot, and tongue. Despite the promising results obtained for distinguishing between the MI tasks of the aforementioned large body-parts, the limited number of control signals produced by these EEG-based BCI systems, which corresponds to the number of MI tasks supported by these systems, can significantly reduce the potential of utilizing such systems to control dextrous prosthetic hands that require a higher number of control signals.

Towards increasing the control's dimensions of the EEG-based BCI systems, few researchers have recently investigated the possibility of decoding the MI tasks related to fine body-parts, such as the MI tasks associated with fingers of the same hand [7], [18], the MI tasks associated with the wrist within the same hand [19], [20], and grasp-related MI tasks that are performed by the same hand [7], [21]. Distinguishing between the EEG signals associated with the MI tasks that are performed by the fine body-parts of the same hand, such as the wrist and fingers, is considered more challenging than distinguishing between the EEG signals associated with the MI tasks performed by different large body-parts, such as the left hand, right hand, feet, and tongue [7], [19], [22], [23]. This can be attributed to several factors, including: (1) The MI tasks related to fine body-parts of the same hand, such as wrist and fingers, activate relatively close and small regions of the primary motor cortex within the same hemisphere of the brain [19], [22], [24]. (2) The limited spatial resolution and the low signal-to-noise ratio (SNR) of the EEG signals degrade the capability of capturing the brain activities generated in response to the MI tasks performed by the wrist and fingers within the same hand [22], [23]. (3) EEG signals are nonstationary signals with time-varying spectral characteristics. This implies that analyzing the EEG signals in the time-domain or

the frequency-domain is considered inadequate [7]. In fact, the nonstationarity nature of the EEG signals imposes the need to represent the EEG signals in a joint time-frequency domain that can better describe the variations in the spectral characteristics of the EEG signals with respect to the time [25].

The aforementioned factors stipulate the necessity of employing robust feature extraction and classification techniques to develop EEG-based BCI systems that can accurately distinguish between the MI tasks performed by fine body-parts of the same hand. Over the past two decades, researchers have utilized various signal processing techniques to extract features from the EEG signals, such as the common spatial patterns (CSP) [26], [27] and Wavelet transform [28]. The obtained handcrafted features were classified into different MI tasks using different classification methods, such as the support vector machines (SVM) [29] classifier. Despite the promising results obtained based on using handcrafted features that are extracted from various signal processing techniques, the process of manually design new features that characterize the most relevant information of the EEG signals is considered challenging [30]. This is due to the difficulty of selecting the most optimal features from a large number of time-domain, frequency-domain, and time-frequency domain features that can be computed based on using different signal processing techniques [30].

Recently, deep learning (DL) methods have achieved promising performance improvements in the fields of image classification and computer vision [31], [32]. The ability of the DL methods to learn latent features from the input images and signals alleviate the need to manually design the features. In this vein, the use of convolutional neural networks (CNNs), which is one of the commonly used DL methods, in image classification problems has lead to a significant improvement in the classification accuracy compared to other approaches that rely on handcrafted features [12], [33]. Despite the fact that the vast majority of the existing MI EEG-based BCI systems are based on using handcrafted features, few researchers have recently explored the possibility of utilizing CNNs to decode MI tasks related to large body-parts [33], such as the left hand, right hand, feet, and tongue. For example, Tabar and Halici [6] utilized the short time Fourier transform (STFT) to convert EEG signals into two-dimensional (2D) images. Then, a CNN is used to extract features from the obtained 2D images. The extracted features are fed into a stack autoencoder network to classify the EEG signals into two classes of MI task, namely the left hand and right hand MI tasks. The average classification accuracy obtained by the proposed approach based on the publicly available BCI competition IV dataset 2b [34] was 74.8%. Lawhern *et al.* [12] proposed a compact CNN (EEGNet) that employs depthwise and spatial convolutions to extract features from the EEG signals. The proposed EEGNet was evaluated using the publicly available BCI competition IV dataset 2a [35], which consists of EEG signals that were recorded for nine healthy subjects while performing four MI tasks related to the left hand, right

hand, feet, and tongue. The average classification accuracy obtained using the EEGNet computed over the four classes of MI tasks was approximately 69%. Dose *et al.* [14] proposed a CNN architecture that consists of two convolutional layers for decoding MI tasks based on EEG signals. In particular, the first layer convolves the EEG signals along the time axis, while the second layer convolves the results of the first layer along the EEG channels. The performance of the proposed CNN architecture was evaluated using the publicly available Physionet EEG Dataset [36], [37], which consists of EEG signals that were recorded for 109 healthy subjects while imagining to opening and closing both feet, both fists, left fist, and right fist. The average classification accuracy obtained using the proposed CNN architecture computed over the four classes of MI tasks was 68.1%. In another study, Schirrmeister *et al.* [13] designed three different CNN architectures, namely shallow, deep, and residual CNN architectures, for decoding MI tasks using EEG signals. The effects of design choices were systematically evaluated for each of the three CNN architectures. Moreover, the classification performance of each architecture was validated against a baseline decoding method, namely the filter bank common spatial patterns (FBCSP). Experimental results, which are obtained based on the publicly available BCI Competition IV dataset 2a [35], show that the shallow and deep architectures were able to archive classification accuracies that are very close to the accuracies obtained using the FBCSP method for decoding MI tasks using EEG signals. In fact, previous studies that investigated the use of DL to decode MI tasks suggest the feasibility of utilizing CNNs to distinguish between the EEG signals associated with MI tasks that are performed by large body-parts, including right hand, left hand, feet, and tongue. However, the existing DL-based studies have not investigated the capability of utilizing CNNs to distinguish between EEG signals that are associated with MI tasks that are performed by fine body-parts of the same hand, such as the wrist and fingers.

In this study, we propose a novel framework that can distinguish between eleven MI tasks that are associated with fine body-parts of the same hand, including the rest, three different grasp-related tasks, the flexion/extension and ulnar/radial deviation tasks of the wrist, and the flexion/extension task of each finger, using EEG signals. In particular, the proposed framework consists of three stages, namely the input, feature extraction, and classification stages. At the input stage, we employ a quadratic time-frequency distribution (QTFD) to analyze the EEG signals in the joint time-frequency domain. The use of a QTFD enables to transform the EEG signals into a set of 2D time-frequency images (TFIs) that describe the distribution of the energy encapsulated within the EEG signals in terms of the time, frequency, and electrode position. At the feature extraction stage, we design a new CNN architecture that can automatically analyze and extract salient features from the TFIs obtained at the input stage. Finally, the features obtained at the second stage are passed to

the classification stage, which in turn assigns each input TFI to one of the eleven MI tasks considered in this work.

To evaluate the performance of our proposed framework, we have recorded EEG signals for eighteen able-bodied subjects and four transradial amputated subjects while performing eleven MI tasks using the same hand. A ten-fold cross-validation procedure is employed to train and test the feature extraction stage and the classification stage of our proposed framework using the TFIs constructed at the input stage from the recorded EEG signals of each subject. The classification performance is computed for each subject in terms of four different evaluation metrics, namely the classification accuracy, recall, precision, and F_1−score. Moreover, the performance of our proposed framework is compared with the performance obtained using handcrafted features that are extracted from the QTFD-based time-frequency representation (TFR) and classified using a multi-class support vector machine (SVM) classifier. The experimental results demonstrate the capability of our proposed framework to decode the eleven MI tasks that are associated with the same hand. In addition, the results obtained using our proposed framework are significantly higher than the results obtained using the QTFD-based handcrafted features which are classified using multi-class SVM classifier. To the best of our knowledge, this is the first study that investigates the possibility of employing CNNs to learn features from a QTFD-based TFR of the EEG signals and to distinguish between eleven MI tasks that are associated with fine body-parts of the same hand for both able-bodied and transradial amputated subjects.

The rest of this paper is structured as follows: Section II describes the experimental protocol and the proposed framework for decoding MI tasks within the same hand. The results and discussion are presented in sections III and IV, respectively. Finally, the conclusion is provided in section V.

II. MATERIALS AND METHODS
A. SUBJECTS

In this study, we utilize an extended version of the EEG dataset described in our previous work [7]. In particular, the EEG signals were recorded for two groups of subjects. The first group of subjects consists of eighteen able-bodied subjects (six females and twelve males, four left-handed and fourteen right-handed, with an average ± standard deviation age of 21.2 ± 2.9 years). We refer to the eighteen subjects in the first group as S_1 to S_{18}. The second group of subjects consists of four male subjects with transradial amputations. The average ± standard deviation age of the amputated subjects is 28.5 ± 6.2 years. Table 1 describes the characteristics of the amputations associated with each of the four amputated subjects. We refer to the four subjects in the second group as A_1 to A_4. Before participating in the experiment, each subject received a thorough explanation of the experimental procedure and signed a consent form. The experimental procedure of this study was approved by the Research Ethics Committee

TABLE 1. Characteristic information about each of the four transradial amputated subjects who participated in this study. For each amputated subject, we provide the dominant hand, amputated hand, elapsed time since amputation, and the percentage of the remaining forearm. The percentage of the remaining forearm is computed as the ratio between the length of the amputated forearm measured from the elbow and the length of the contralateral intact forearm measured from the elbow to the wrist [38]. The right hand and left hand are abbreviated as RH and LH, respectively.

Subject	Dominant hand	Amputated hand	Elapsed time since amputation (years)	Remaining percentage of the forearm
A_1	RH	LH	3.5	28%
A_2	RH	RH	1.5	10%
A_3	LH	LH	4	95%
A_4	RH	RH	5	37%

at the German Jordanian University and was conducted in accordance with the Declaration of Helsinki.

B. EXPERIMENTAL PROCEDURE

At the beginning of the experiment, each subject was asked to sit on a chair and to relax her/his arms on a table located in front of her/him. A computer screen was placed on the table at a distance of approximately 50 cm from the subject and employed to display various visual cues. In particular, each visual cue notifies the subject to imagine performing a specific hand task. The MI tasks considered in this work can be grouped into four categories: (1) Rest (M_1), (2) Grasp-related tasks, including the small diameter grasp (M_2), lateral grasp (M_3), and extension-type grasp (M_4), (3) Wrist-related tasks, including the ulnar and radial deviation (U/R) of the wrist (M_5) and the flexion and extension (F/E) of the wrist (M_6), and (4) Fingers-related tasks, including the F/E of the index finger (M_7), F/E of the middle finger (M_8), F/E of the ring finger (M_9), F/E of the little finger (M_{10}), and F/E of the thumb finger (M_{11}). Figure 1 shows sample images of the visual cues associated with each hand imagery task.

(a) Rest task

(b) Grasp-related tasks
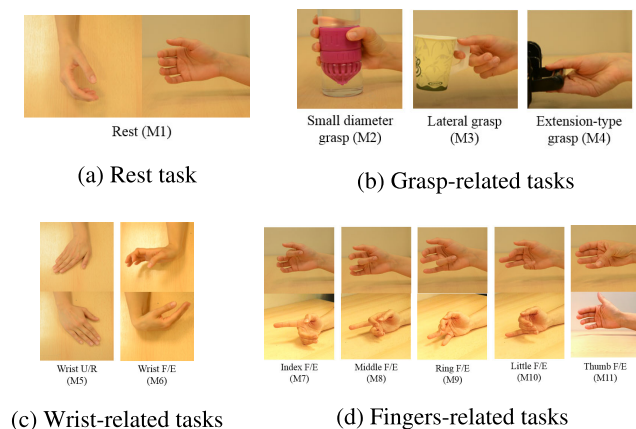
(c) Wrist-related tasks

(d) Fingers-related tasks

FIGURE 1. Sample images of the visual cues employed for each MI task [7].

The experiment starts by displaying a visual cue on the computer screen for three seconds. Then, a black screen is displayed to prompt the subject to start imagining to perform the task associated with the displayed visual cue. The able-bodied subjects were asked to imagine performing each task using their right hands, while the transradial amputated subjects were asked to imagine performing the tasks using their

lost hands. Moreover, the subjects were asked to retain their eyes closed during the imagination of each hand task. The number of trials recorded per each imaginary hand task for each able-bodied subject is 40 trials, while the number of trials recorded per each imaginary hand task for each transradial amputated subject is 56 trials. The lengths of the recorded trials vary depending on the hand task being imagined. In particular, the duration of each trial, including the duration of the visual cue, recorded during the rest, wrist-related tasks, and fingers-related tasks is equal to 10 s. On the other hand, the duration of each trial, including the duration of the visual cue, associated with the small diameter grasp, lateral grasp, and extension-type grasp tasks is equal to 14 s, 14 s, and 12 s, respectively.

C. RECORDING AND PREPROCESSING OF EEG SIGNALS

The BioSemi ActiveTwo EEG system (Biosemi B.V., Amsterdam, Netherlands) was used to record the EEG signals using 16 Ag/AgCl electrodes at a sampling rate of 2048 Hz. The utilized EEG electrodes are arranged on the scalp according to the $10 - 20$ international electrode placement system at the following locations: C_3, C_4, Fp_1, Fp_2, F_4, F_z, F_3, T_7, C_z, T_8, P_4, P_z, P_3, O_1, O_z, and O_2. The employed 16 EEG electrodes are referenced to the common mode sense (CMS)/ driven right leg (DRL) at the C_1 and C_2 locations. The recorded EEG signals were filtered by applying a bandpass filter with a bandwidth of $0.5 - 32.5$ Hz [21], [39]. The filtered EEG signals were downsampled to 256 Hz. Moreover, the automatic artifact removal (AAR) toolbox [40], [41] was employed to reduce the muscle and electrooculography (EOG) artifacts in the filtered EEG signals.

D. DECODING MI TASKS OF THE SAME HAND

In this study, we propose a novel framework to decode the eleven MI tasks of the same hand that are described in subsection II-B. In particular, the proposed framework consists of three stages, namely the input, feature extraction, and classification stages. At the input stage, we employ a quadratic time-frequency distribution (QTFD) to analyze the EEG signals in the joint time-frequency domain. The use of a QTFD enables to transform the EEG signals into a set of 2D time-frequency images (TFIs) that describe the distribution of the energy encapsulated within the EEG signals in terms of the time, frequency, and electrode position. Detailed description of the input stage is provided in subsection II-D1. At the
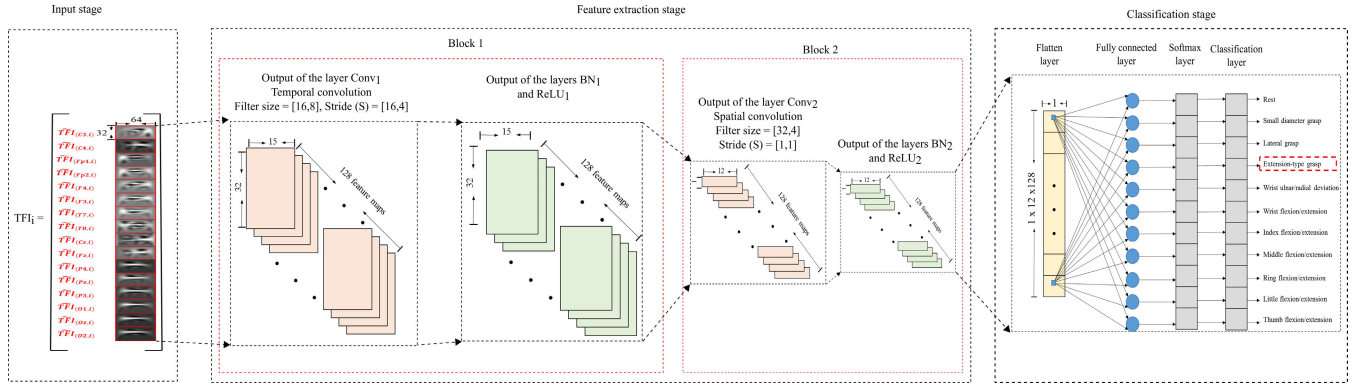
FIGURE 2. The architecture of the proposed framework for decoding MI tasks of the same hand.

feature extraction stage, we construct a CNN to automatically analyze and extract salient features from the TFIs obtained at the input stage. The structure of the CNN constructed at the feature extraction stage is provided in subsection II-D2. Finally, the features obtained at the second stage are passed to the classification stage, which assigns each input TFI to one of the eleven MI tasks considered in this work. The structure of the classification stage is provided in subsection II-D3. Figure 2 illustrates the architecture of our proposed framework for decoding MI tasks of the same hand.

1) INPUT STAGE

EEG signals are nonstationary signals that comprise time-varying spectral components [7]. Hence, analyzing the EEG signals in the time-domain or the frequency-domain may not capture the spectral variations of the EEG signals over time [42]. In this study, we propose the use of a QTFD, namely the Choi-Willimas Distribution (CWD) [43], to analyze the EEG signals. The CWD can be viewed as a two-dimensional (2D) transformation that maps the original time-domain EEG signals into a joint time-frequency domain which has an excellent resolution in both the time and frequency domains [7], [25], [43]–[45]. In particular, the use of the CWD enables the conversion of the EEG signals into 2D TFIs that can quantify the distribution of the energy encapsulated in the EEG signals over the time and frequency domains [25]. Therefore, the objective of the input stage is to transform the EEG signals into TFIs.

In order to compute the TFIs of the EEG signals, we utilize a sliding window that divides the EEG signal of each electrode into a set of overlapped segments, such that the size of each segment is 64 samples and the overlap between any two consecutive segments is 10 samples. For a given EEG segment, denoted by $s_{(c,i)}(t)$, where $c \in E = \{C_3, C_4, Fp_1, Fp_2, F_4, F_z, F_3, T_7, C_z, T_8, P_4, P_z, P_3, O_1, O_z, O_2\}$ represents the EEG electrode associated with the EEG segment and i represents the time index of the first sample in the window position that corresponds to the segment, the CWD-based TFI of $s_{(c,i)}(t)$, denoted as $TFI_{(c,i)}$, can be constructed as follows [25], [44]:

1. Compute the analytic signal of $s_{(c,i)}(t)$, denoted as $x_{(c,i)}(t)$, as expressed below:

$$x_{(c,i)}(t) = s_{(c,i)}(t) + j\mathcal{H}\{s_{(c,i)}(t)\}, \quad (1)$$

where $\mathcal{H}\{\cdot\}$ is the Hilbert transform [46] and $j = \sqrt{-1}$.

2. Compute the CWD of $x_{(c,i)}(t)$, denoted as $TFI_{(c,i)}(t, f)$, as follows [43], [47]:

$$TFI_{(c,i)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \chi_{x_{(c,i)}}(\mu, \nu) \kappa(\mu, \nu) e^{-j2\pi(f\nu + t\mu)} \partial_\nu \partial_\mu, \quad (2)$$

where $\chi_{x_{(c,i)}}(\mu, \nu)$ is the ambiguity function of $x_{(c,i)}(t)$, and $\kappa(\mu, \nu)$ is a time-frequency smoothing kernel. In particular, $\chi_s(\mu, \nu)$ represents the Fourier transform of the auto-correlation function of $x_{(c,i)}(t)$ and can be computed as [43], [47]:

$$\chi_{x_{(c,i)}}(\mu, \nu) = \int_{-\infty}^{\infty} x_{(c,i)}\left(t + \frac{\nu}{2}\right) x_{(c,i)}^*\left(t - \frac{\nu}{2}\right) e^{j2\pi\mu t} \partial t, \quad (3)$$

where $x_{(c,i)}^*(\cdot)$ is the complex conjugate of $x_{(c,i)}(\cdot)$. The time-frequency smoothing kernel, $\kappa(\mu, \nu)$, can be expressed as follows [43]:

$$\kappa(\mu, \nu) = \exp\left(-\frac{\mu^2 \nu^2}{\omega^2}\right), \quad (4)$$

where $\omega > 0$ is a smoothing parameter with an experimentally-selected value of 0.5.

The size of the image $TFI_{(c,i)}$ is equal to 128×64, where 128 and 64 represent the number of samples along the frequency and time axes, respectively. As shown in Fig. 3, at each window position, we construct 16 CWD-based TFIs, where each image corresponds to an EEG segment that is associated with one of the 16 EEG electrodes defined in the set E. After that, the CWD-based TFIs obtained at each window position are cropped to keep the frequency components within the range $0.5 - 32.5$ Hz for each CWD-based TFI. We denote each of the cropped TFIs at a specific window position as $\widetilde{TFI}_{(c,i)}$. The size of the image $\widetilde{TFI}_{(c,i)}$ is $N_f \times N_t$, where $N_f = 32$ and $N_t = 64$ represent the number of frequency and time samples, respectively. Each image $\widetilde{TFI}_{(c,i)}$ can be divided
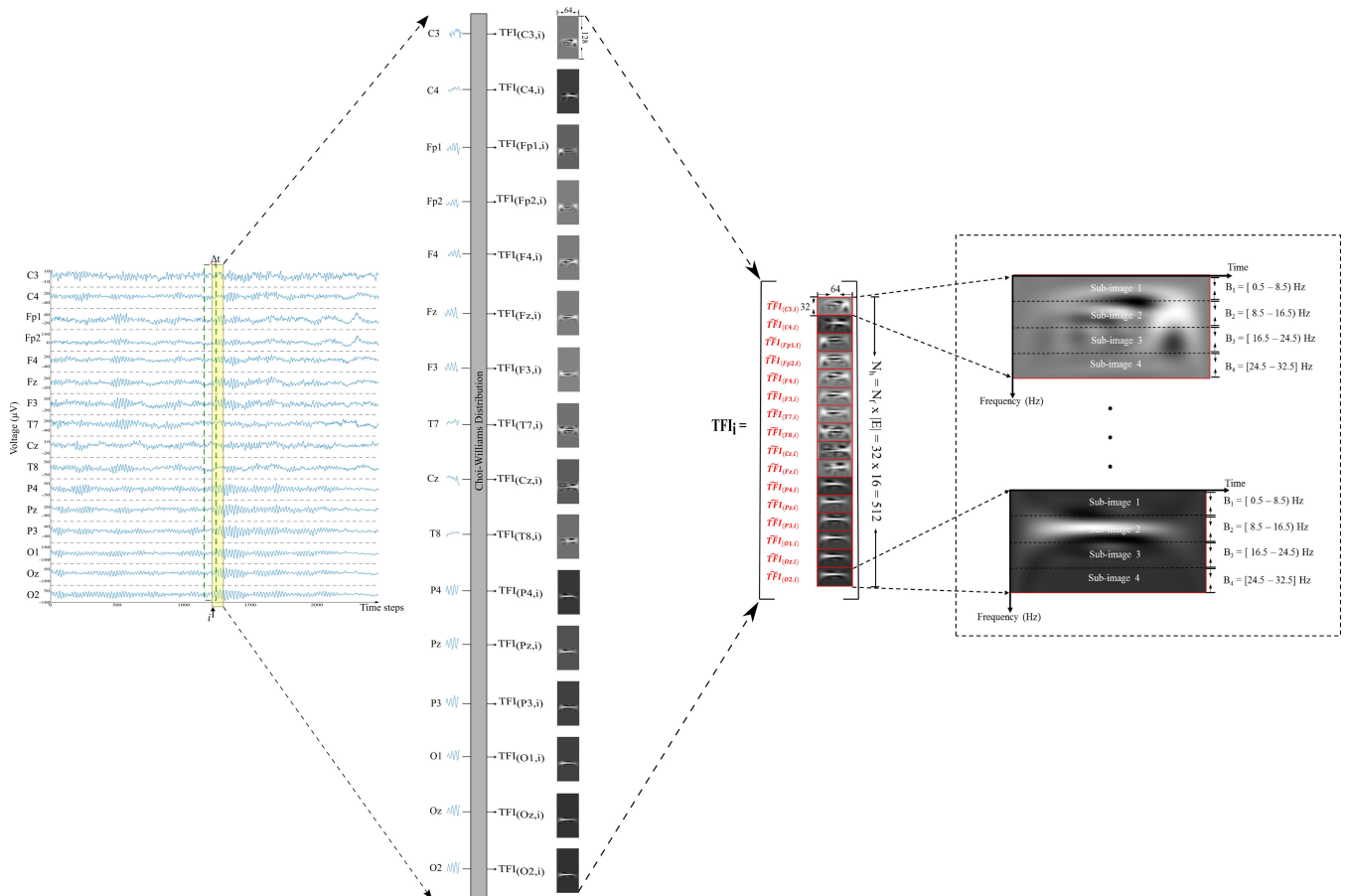
FIGURE 3. Structure diagram that demonstrates the procedure employed to construct an input image for the feature extraction stage, namely TFI_i, using the CWD-based TFIs obtained for the EEG segments at window position i, namely the images $\widetilde{TFI}_{(c,i)}$ for all $c \in E$. The yellow rectangle represents the current position of the sliding window, the green dashed rectangle represents the previous position of the sliding window, and $\triangle t$ represents the overlap between two consecutive window positions and is equal to 10 samples.

along the frequency dimension into four sub-images, namely *sub-image₁*, *sub-image₂*, *sub-image₃*, and *sub-image₄*, where each sub-image is associated with one of four frequency bands, detonated by B_1, B_2, B_3, and B_4, that are described in Table 2. Finally, at each window position, the images $\widetilde{TFI}_{(c,i)}$ for all $c \in E$ are vertically combined to form a new image, denoted as TFI_i. The size of the constructed TFI_i is $N_h \times N_t$, where $N_h = N_f \times |E|$, $|E|$ is the number of EEG electrodes. Figure 3 describes the procedure used to construct the TFI_i using the EEG signals that were recorded for the first able-bodied subject, namely S_1, while imagining to perform the extension-type grasp task. In addition, Fig. 3 illustrates the sub-images within each of the 16 images $\widetilde{TFI}_{(c,i)}$ for all $c \in E$ of the constructed TFI_i.

2) FEATURE EXTRACTION STAGE

The input stage generates TFIs that describe the distribution of the energy of the EEG signals encapsulated within each window position in terms of the time, frequency, and electrode position. For example, the image TFI_i comprises the time-frequency distribution of the energy encapsulated within the EEG signals at the ith window position for each of the

16 EEG electrodes. Figure 4 provides samples of the TFIs constructed at the ith window position of the EEG signals recorded for the subject S_1 while imagining to perform the extension-type grasp, the F/E of the middle finger, the F/E of the wrist, and the rest tasks, respectively. In addition, Fig. 4 shows that the energy distribution of the EEG signal associated with each EEG electrode is varying over time. Moreover, for each TFI_i, the TFIs associated with the 16 EEG electrodes, namely $\widetilde{TFI}_{(c,i)}$ for all $c \in E$, exhibit large variations in the energy distribution over time. This imposes the requirement to analyze the spectral variations occurring over time within the EEG signal of each electrode as well

TABLE 2. The frequency bands associated with the four sub-images within the image $\widetilde{TFI}_{(c,i)}$.

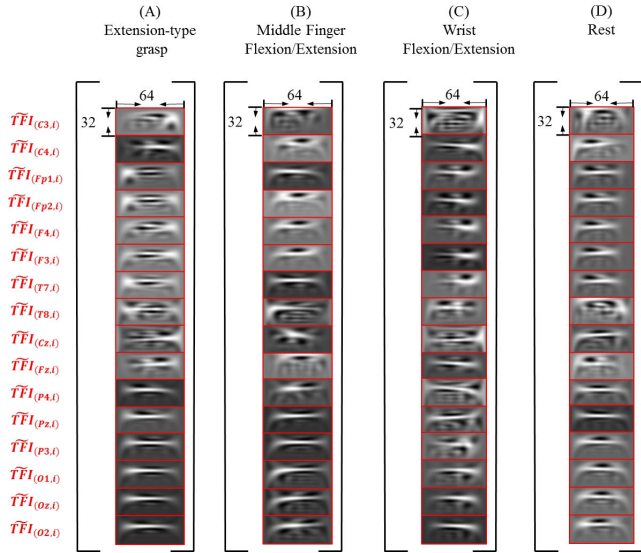| Sub-image	Frequency band (Hz)
$Sub\text{-}image_1$	$B_1 = [0.5 - 8.5)$
$Sub\text{-}image_2$	$B_2 = [8.5 - 16.5)$
$Sub\text{-}image_3$	$B_3 = [16.5 - 24.5)$
$Sub\text{-}image_4$	$B_4 = [24.5 - 32.5]$

FIGURE 4. Samples of the TFIs constructed at the ith window position of the EEG signals recorded for the subject S_1 while imagining to perform (A) the extension-type grasp, (B) the F/E of the middle finger, (C) the F/E of the wrist, and (D) the rest task, respectively.

as across the EEG signals of different electrodes. Therefore, the feature extraction stage aims to learn salient features from each TFI_i that can be used to distinguish between different MI tasks within the same hand.

The feature extraction stage of our proposed framework is realized using a CNN that consists of two blocks of layers as shown in Fig. 2. The first block, namely block 1, consists of three layers, including the convolutional layer ($Conv_1$), batch normalization layer (BN_1), and rectified linear unit layer ($ReLU_1$). The $Conv_1$ layer consists of a number of neurons that are connected to local regions of the input image TFI_i. The spatial extent of these local regions is determined based on the size of the filters employed to scan the input image, where each filter consists of a set of weights and one bias parameter that are learned during the training of the CNN. The size of the filters in the $Conv_1$ layer, denoted by $F_h^{Conv_1} \times F_w^{Conv_1}$, where $F_h^{Conv_1}$ and $F_w^{Conv_1}$ represent the height and width of each filter, respectively. To scan the input image TFI_i, the filters in the $Conv_1$ layer are shifted along the horizontal and vertical directions of the image TFI_i. The amount of the shift along the horizontal and vertical directions is determined based on the value of the stride parameter, denoted as $S_{Conv_1} = [S_v, S_h]$, where S_v and S_h determine the value of the vertical and horizontal shifts of the filters, respectively.

In this work, the height and width of the filters in the $Conv_1$ layer are selected experimentally to be $F_h^{Conv_1} = 16$ and $F_w^{Conv_1} = 8$. In addition, the stride parameter is set to $S_{Conv_1} = [16, 4]$. The number of filters employed in the $Conv_1$ layer is selected experimentally to be 128 filters. The selected size of the filters along with the selected value of the stride parameter of the $Conv_1$ layer enable to analyze the sub-images within each of the 16 images $\widetilde{TFI}_{(c,i)}$ for all $c \in E$ of the input image TFI_i. In particular, the $Conv_1$ layer

can be viewed as a time-frequency convolutional layer that convolves the first two sub-images, which correspond to the frequency bands B_1 and B_2, and the second two sub-images, which correspond to the frequency bands B_3 and B_4, within the image $\widetilde{TFI}_{(c,i)}$ along the time axis to produce a set of feature maps. Specifically, the feature map that is generated using the rth filter of the $Conv_1$ layer, denoted as $fm_{(r,Conv_1)}$, can be computed as follows:

$$fm_{(r,Conv_1)} = W_{(r,Conv_1)} \otimes TFI_i + b_{(r,Conv_1)}, \quad (5)$$

where $\otimes$ represents the convolution operator, and $W_{(r,Conv_1)}$ and $b_{(r,Conv_1)}$ represent the weights matrix and the bias parameter associated with the rth filter of the $Conv_1$ layer, respectively. The number of feature maps generated at the output of the $Conv_1$ layer is equal to 128 maps, which is equal to the number of filters employed in the $Conv_1$ layer, and the size of each feature map is 32×15. Figure 5 illustrates the time-frequency convolution carried out at the $Conv_1$ layer for an input image TFI_i.

The feature maps obtained at the output of the $Conv_1$ layer are passed on to the next layer in block 1, which is the BN_1 layer. The BN_1 layer normalizes the mean and variance of the features during the training of the CNN. This in turn can accelerate the training of the CNN and reduce the occurrence of overfitting [13]. The size of each feature map obtained at the output of the BN_1 layer is 32×15. After that, the normalized feature maps are passed on to the next layer in block 1, which is the $ReLU_1$ layer. The $ReLU_1$ layer can be viewed as an activation function that applies a threshold operation to each feature within each feature map. In particular, the $ReLU_1$ layer sets all the features within a particular feature map that have negative values to zero. Such an activation function introduces sparsity and nonlinearity to the CNN structure that can reduce the effect of the small variations in the input images [6], [13], [48]. The size of each feature map obtained at the output of the $ReLU_1$ layer is 32×15. The structure of the feature maps obtained at the output of the $ReLU_1$ layer is illustrated in Fig. 6. Specifically, within each feature map, the features extracted from the images $\widetilde{TFI}_{(c,i)} \ \forall c \in E$ are arranged in 16 sub-maps each of size 2×15. These sub-maps are arranged according to the same order of the EEG electrodes in the set E, where the first sub-map contains the features that are associated with the EEG electrode C_3 and the last sub-map contains the features that are associated with the EEG electrode O_2. Therefore, the feature maps obtained at the output of block 1 comprise features that characterize the variations of the energy distribution within each individual EEG electrode without taking into consideration the variations of the energy distribution across different EEG electrodes.

In order to characterize the variations of the energy distribution across different EEG electrodes, we fed the feature maps obtained at the output of block 1 to the second block in the feature extraction stage, namely block 2, which consists of three layers, including the convolutional layer ($Conv_2$), batch normalization layer (BN_2), and rectified linear unit
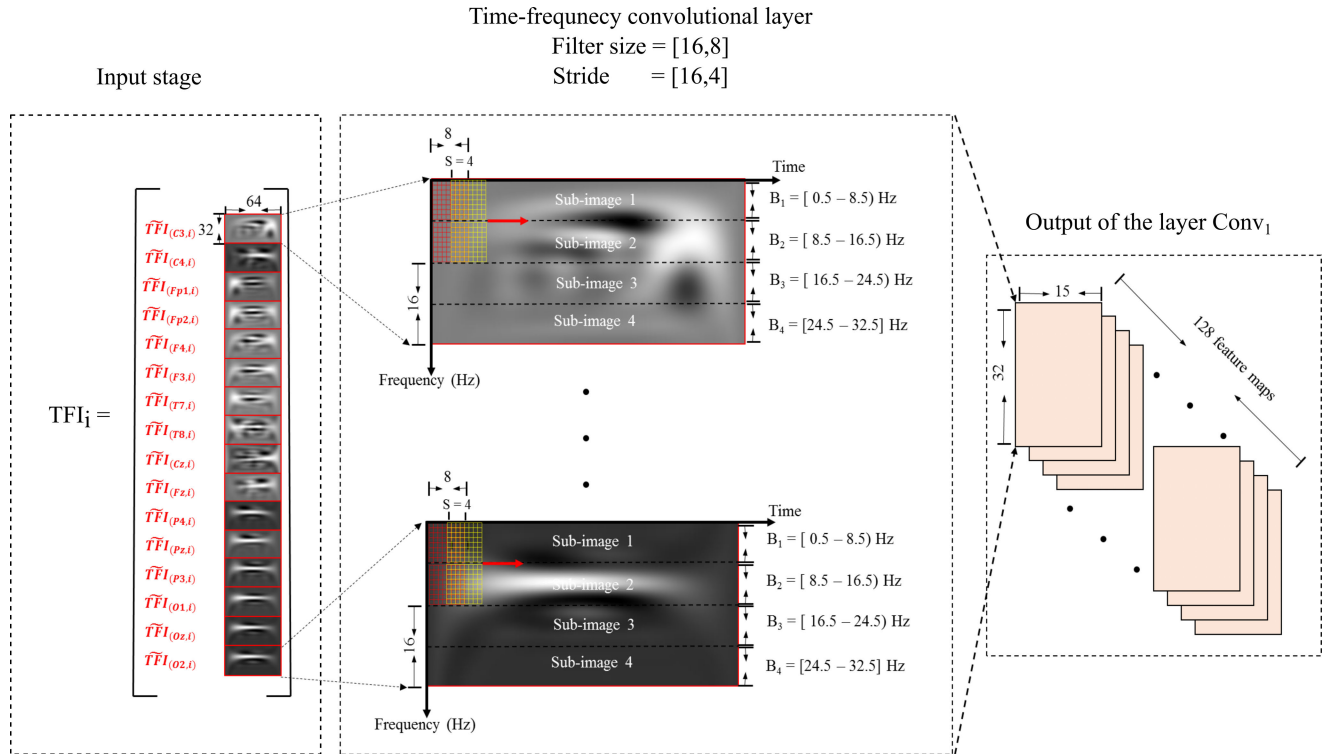
FIGURE 5. Graphical illustration of the time-frequency convolution performed at the $Conv_1$ layer. The TFI associated with each EEG electrode is divided into four sub-images, where each sub-image is associated with a frequency band. The red grid represents the current position of a particular filter, while the yellow grid represents the next position of the filter.
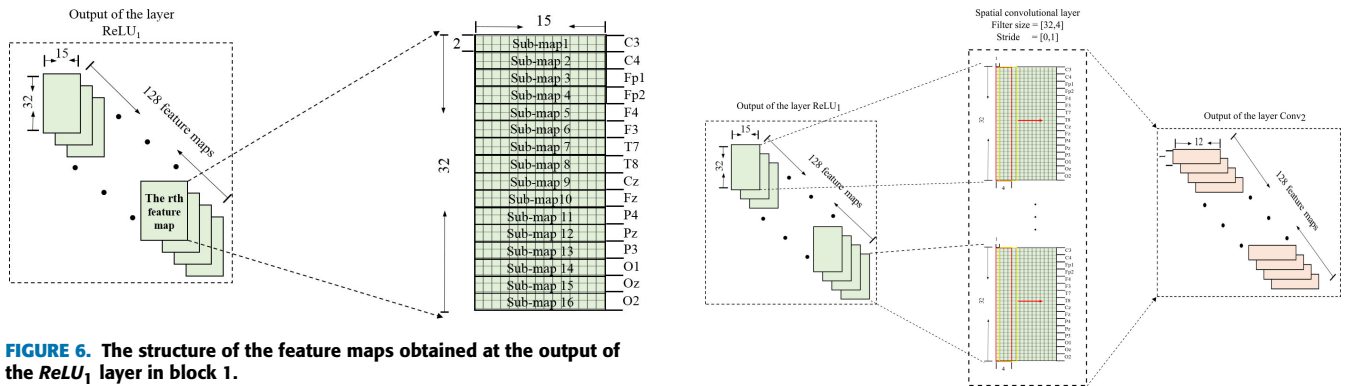

FIGURE 6. The structure of the feature maps obtained at the output of the $ReLU_1$ layer in block 1.

FIGURE 7. Graphical illustration of the spatial convolution performed at the $Conv_2$ layer. The red rectangle represents the current position of a particular filter, while the yellow rectangle represents the next position of the filter.

layer ($ReLU_2$). The $Conv_2$ layer consists of a number of neurons that are connected to local regions of the feature maps obtained at the output of the $ReLU_1$ layer in block 1. Similar to the $Conv_1$ layer, the spatial extent of the local regions defined at the $Conv_2$ layer is determined based on the size of the filters employed to scan the feature maps, where each filter consists of a set of weights and one bias parameter that are learned during the training of the CNN. The size of each filter is set to $F_h^{Conv_2} \times F_w^{Conv_2}$, where $F_h^{Conv_2} = 32$ and $F_w^{Conv_1} = 4$ represent the height and width, respectively, of each filter. In addition, the stride of the $Conv_2$ layer, denoted by $S_{Conv_2} = [S_v, S_h]$, is set to $S_{Conv_2} = [0, 1]$. The number of filters employed in the $Conv_2$ layer is equal to 128 filters. These selected values of the filters sizes and the stride parameter of the $Conv_2$ layer enable the scanning of the

features associated with all EEG electrodes along the time axis. Therefore, the $Conv_2$ layer can be viewed as a spatial convolutional layer that convolves all the sub-maps within a particular feature map along the time axis to produce a new set of feature maps. The number of feature maps generated at the output of the $Conv_2$ layer is equal to 128 maps, which is equal to the number of filters employed in the $Conv_2$ layer, and the size of each feature map is 1×12. Figure 7 illustrates the spatial convolution performed at the $Conv_2$ layer for each of the feature maps obtained from the $ReLU_1$ layer.

Similar to block 1, the feature maps obtained at the output of the $Conv_2$ layer are passed on to the BN_2 layer

for normalization. The size of each feature map obtained at the output of the BN_2 layer is 1×12. Finally, the normalized feature maps are passed through the $ReLU_2$ layer. The size of each feature map obtained at the output of the $ReLU_2$ layer is 1×12. The feature maps obtained at the output of the $ReLU_2$ layer represent the features extracted from the input image TFI_i. These feature maps are passed on to the classification stage in order to determine the class of the MI task associated with the input image TFI_i.

3) CLASSIFICATION STAGE

The objective of the classification stage is to assign each TFI_i, which is obtained at the input stage, to one of the eleven MI tasks described in subsection II-B using the features learned at the feature extraction stage. In particular, the classification stage consists of four layers, including flatten layer, fully connected layer, softmax layer, and classification layer, as shown in Fig. 2. At the flatten layer, the feature maps obtained at the output of the $ReLU_2$ layer in the feature extraction stage, which are computed for a particular TFI_i, are reshaped into one-dimensional (1D) feature vector. The size of the 1D feature vector obtained at the output of the flatten layer is $R \times 1$, where R equals to the number of feature maps obtained at the output of the $ReLU_2$ layer multiplied by the size of each feature map. The 1D feature vector constructed at the output of the flatten layer is passed on to the next layer in the classification stage, which is the fully connected layer. The fully connected layer consists of eleven neurons, where each neuron is connected to all the features in the 1D feature vector. Specifically, the fully connected layer combines all the features learned via the feature extraction stage to classify the input image TFI_i. The outputs of the fully connected layer, which are obtained by multiplying the 1D feature vector by a trainable weight matrix and then add a trainable bias vector, are passed on to the softmax layer. The softmax layer normalizes the outputs of the fully connected layer to ensure that the output values of the softmax layer are positive numbers that sum to one. Therefore, each of the eleven outputs of the softmax layer can be viewed as the probability value that the input image TFI_i belongs to one of the eleven MI tasks. Finally, the eleven probability values obtained at the output of the softmax layer are passed on to the classification layer, which assigns the input image TFI_i to the MI task that has the highest probability value.

4) TRAINING AND TESTING OF THE PROPOSED FRAMEWORK

For each subject, we utilize a ten-fold cross-validation procedure to construct the framework described in subsection II-D. In particular, the EEG signals associated with each trial of a particular subject are converted into a set of labeled TFIs, where the label of each TFI represents the MI task associated with the trial from which the TFI was extracted. The construction of each TFI is based on the procedure employed at the input stage, which is described in subsection II-D1. Concretely, the set of labeled TFIs extracted from the jth trial,

denoted as T_j, of the kth MI task, denoted as M_k, of the rth subject is expressed as follows:

$$D^{(j,k,r)} = \{(TFI_1^{T_j}, M_k), \cdots, (TFI_i^{T_j}, M_k), \cdots, (TFI_l^{T_j}, M_k)\}, \quad (6)$$

where $1 \leq k \leq 11$, $1 \leq i \leq l$, l, and r represent the index of the MI task, the window position, the number of window positions in the trial Tj, and the index of the subject, respectively. The pair $(TFI_i^{T_j}, M_k)$ represents the TFI extracted from the EEG segments located at window position i within the trial T_j of the MI task M_k. The labeled data obtained from all trials of all MI tasks of a particular subject are divided randomly into ten folds. The feature extraction stage and classification stage of the proposed framework are trained using nine folds and tested using the remaining fold. The ten-fold cross-validation procedure is repeated for ten times and the average classification performance for each subject is computed over the ten repetitions. During each repetition of the ten-fold cross-validation procedure, the parameters of the feature extraction stage and classification stage in our proposed framework, including the weights and biases of the filters in the $Conv_1$ and $Conv_2$ layers as well as the weight matrix and bias vector of the neurons in the fully connected layer, are learned using a mini-batch stochastic gradient decent (SGD) optimization method that minimizes the categorical cross-entropy loss function using the back-propagation algorithm [31]. In particular, the training process was run for 40 epochs, where each epoch represents one pass over all the training samples. In addition, the mini-batch size, which is the number of training samples used during each iteration of the training process, is selected to be 64 samples. The learning rate of the SGD optimization method is selected to be 0.01. Finally, the hyperparameters of the feature extraction stage and classification stage in our proposed framework, including the number of filters, the size of the filters, and the stride of each convolutional layer, were experimentally selected [6], [12]–[14], [48] as described in section II-D2.

E. PERFORMANCE EVALUATION

The proposed framework is implemented and run on a workstation with an Intel Xeon E5-1620 3.50GHz 4 cores CPU, 8 GB RAM, and Nvidia Quadro K620 GPU. To quantify the classification performance of the constructed framework, which is measured based on the testing folds, we employ five standard evaluation metrics, namely the classification accuracy (CA), recall (REC), precision (PRE), and F_1-score (F_1), that are computed as follows [49]:

$$CA = \frac{(tp + tn)}{(tp + tn + fp + fn)} \times 100\%, \quad (7)$$

$$REC = \frac{tp}{(tp + fn)} \times 100\%, \quad (8)$$

$$PRE = \frac{tp}{(tp + fp)} \times 100\%, \quad (9)$$

$$F_1 = 2 \times \frac{(REC \times PRE)}{(REC + PRE)} \times 100\%, \quad (10)$$

where *tp*, *tn*, *fp*, and *fn* represent the number of true positive, true negative, false positive, and false negative cases, respectively.

The runtime of the proposed framework is quantified by computing the average ± standard deviation values of three metrics, namely the training time, construction time, and classification time, for the able-bodied subjects and transradial amputated subjects. In particular, the training time is defined as the time required to complete one of the ten repetitions of the cross-validation procedure employed to train our proposed framework. The construction time is defined as the time required to construct the image TFI_i at the input stage. The classification time is defined as the time required to classify the image TFI_i at the classification stage into one of the eleven MI tasks.

Finally, we compare the classification performance of our proposed framework with the classification performance obtained using conventional CWD-based handcrafted features and a multi-class SVM classifier.

III. EXPERIMENTAL RESULTS

In this section, we present the results of our proposed framework obtained based on the EEG signals of the able-bodied subjects and transradial amputated subjects. Moreover, we provide the runtime analysis of our proposed framework. Finally, we compare the performance of our proposed framework with the performance obtained using conventional handcrafted features that are extracted from the CWD-based TFR of the EEG signals and classified using a multi-class SVM classifier.

A. RESULTS OF THE ABLE-BODIED SUBJECTS

In this section, we present the classification results of our proposed framework obtained based on the EEG signals of the able-bodied subjects. Figure 8 shows the CA values and the corresponding standard deviation (STD) values obtained for each MI tasks per each able-bodied subject. In particular, Fig. 8a shows the CA values and the corresponding STD values computed for the rest MI task, namely M_1, per each able-bodied subject. In fat, the average ± STD CA value computed over all able-bodied subjects for the M_1 task is $72.9\% \pm 1.1\%$. Figure 8b shows the CA values and the corresponding STD values computed for the grasp-related MI tasks, namely M_2, M_3, and M_4, per each able-bodied subject. The average ± STD CA values computed over all able-bodied subjects for the M_2, M_3, and M_4 tasks are $79.7\% \pm 0.6\%$, $81.8\% \pm 0.5\%$, and $76.1\% \pm 0.7\%$, respectively. Figure 8c presents the CA values and the corresponding STD values computed for the wrist-related MI tasks, namely M_5 and M_6, per each able-bodied subject. The average ± STD CA values computed over all able-bodied subjects for the M_5 and M_6 tasks are $70.9\% \pm 0.7\%$ and $69.5\% \pm 0.9\%$, respectively. Figure 8d shows the CA values and the corresponding STD values computed for the finger-related MI tasks, namely M_7, M_8, M_9, M_{10}, and M_{11}, per each able-bodied subject. The average ± STD CA values computed over all able-bodied

subjects for the M_7, M_8, M_9, M_{10}, and M_{11} tasks are $70.9\% \pm 0.7\%$, $70.8\% \pm 0.8\%$, $71.8\% \pm 0.8\%$, $68.3\% \pm 1.1\%$, and $68.2\% \pm 0.6\%$, respectively. Finally, Fig. 8e presents the CA values and the corresponding STD values computed over all MI tasks per each able-bodied subject. The average ± STD CA value computed over all MI tasks for all able-bodied subjects is $73.7\% \pm 0.5\%$.

Table 3 shows the values of the REC, PRE, and F_1 metrics computed for each MI task per each able-bodied subject. Moreover, the last four rows in Table 3, which are associated with the columns labeled as S_1 to S_{18}, provide the average values of the REC, PRE, and F_1 metrics computed over all MI tasks per each able-bodied subject. In addition, the last column in Table 3 provides the average values of the REC, PRE, and F_1 metrics computed over all able-bodied subjects for each MI task. Finally, the last four rows in the last column of Table 3 provide the average values of the REC, PRE, and F_1 metrics computed over all MI tasks for all able-bodied subjects. In particular, the overall average values of the REC, PRE, and F_1 metrics computed over all able-bodied subjects and across all MI tasks are 72.8%, 73.2%, and 72.9%, respectively.

B. RESULTS OF THE TRANSRADIAL AMPUTATED SUBJECTS

In this section, we present the classification results of our proposed framework obtained based on the EEG signals of the transradial amputated subjects. Figure 9 presents the CA values and the corresponding STD values obtained for each MI tasks per each transradial amputated subject. In particular, Fig. 9a shows the CA values and the corresponding STD values computed for the rest MI task, namely M_1, per each transradial amputated subject. In fact, the average ± STD CA value computed over all transradial amputated subjects for the M_1 task is $81.8\% \pm 0.5\%$. Figure 9b shows the CA values and the corresponding STD values computed for the grasp-related MI tasks, namely M_2, M_3, and M_4, per each transradial amputated subject. The average ± STD CA values computed over all transradial amputated subjects for the M_2, M_3, and M_4 tasks are $79.1\% \pm 0.5\%$, $78.1\% \pm 0.6\%$, and $75.2\% \pm 1.3\%$, respectively. Figure 9c shows the CA values and the corresponding STD values computed for the wrist-related MI tasks, namely M_5 and M_6, per each transradial amputated subject. The average ± STD CA values computed over all transradial amputated subjects for the M_5 and M_6 tasks are $71.1\% \pm 0.5\%$ and $66.7\% \pm 0.3\%$, respectively. Figure 9d shows the CA values and the corresponding STD values computed for the finger-related MI tasks, namely M_7, M_8, M_9, M_{10}, and M_{11}, per each transradial amputated subject. The average ± STD CA values computed over all transradial amputated subjects for the M_7, M_8, M_9, M_{10}, and M_{11} tasks are $71.3\% \pm 0.9\%$, $68.9\% \pm 0.2\%$, $66.2\% \pm 0.4\%$, $62.6\% \pm 0.8\%$, and $72.2\% \pm 0.3\%$, respectively. Finally, Fig. 9e shows the CA values and the corresponding STD values computed over all MI tasks per each transradial

TABLE 3. The results obtained for each able-bodied subject expressed in terms of the values of the REC, PRE, and F_1 evaluation metrics computed for each MI task.

Class	Metric	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	S_{18}	Average across subjects
M_1	REC	97.8	66.2	47.5	69.2	79.1	75.9	82.7	66.3	69.2	68.4	75.5	89.4	75.5	76.6	99.7	58.0	67.4	48.2	72.9
	PRE	96.9	66.9	48.1	69.3	80.0	73.9	88.6	64.9	67.9	69.9	66.7	88.9	71.4	76.5	98.0	61.0	62.5	51.7	72.4
	F_1	97.3	66.5	47.8	69.2	79.5	74.9	85.5	65.6	68.6	69.0	70.6	89.2	73.4	76.5	98.8	59.3	64.9	49.8	72.6
M_2	REC	99.0	80.2	60.3	70.4	82.1	74.6	96.6	61.8	75.2	70.0	88.9	88.8	85.6	96.2	96.7	70.9	66.3	71.0	79.7
	PRE	98.2	74.7	57.9	66.1	75.3	71.8	94.5	60.8	70.3	65.7	88.7	86.9	82.1	91.8	95.6	67.7	63.1	65.3	76.5
	F_1	98.6	77.3	59.0	68.1	78.5	73.2	95.5	61.2	72.6	67.8	88.8	87.8	83.8	94.0	96.1	69.3	64.6	68.0	78.0
M_3	REC	99.3	82.8	69.5	71.2	80.2	73.9	96.3	71.5	67.8	70.7	90.2	93.5	85.3	95.2	94.6	84.4	71.1	75.7	81.8
	PRE	97.3	76.2	66.3	66.1	76.3	66.3	94.7	68.1	69.6	67.0	88.6	90.8	80.5	91.0	93.0	79.5	68.6	71.8	78.4
	F_1	98.3	79.3	67.8	68.5	78.2	69.8	95.5	69.7	68.7	68.8	89.4	92.1	82.8	93.1	93.8	81.9	69.8	73.7	80.1
M_4	REC	97.1	69.1	65.1	60.5	73.6	67.3	97.8	62.7	63.9	83.3	82.7	95.7	77.3	87.6	91.5	77.9	59.8	56.2	76.1
	PRE	97.5	68.6	64.8	60.6	71.3	68.6	95.4	61.9	64.7	79.5	81.6	91.5	78.5	85.5	92.6	76.8	57.8	56.9	75.2
	F_1	97.3	68.8	64.9	60.4	72.4	68.0	96.5	62.3	64.3	81.3	82.1	93.5	77.9	86.5	92.0	77.2	58.8	56.5	75.6
M_5	REC	97.9	69.9	51.4	48.6	67.0	62.0	96.0	65.4	57.9	58.8	85.8	85.1	71.4	85.7	94.1	57.8	56.0	66.0	70.9
	PRE	96.7	71.6	56.2	48.8	72.2	70.2	94.7	65.6	57.1	62.1	86.6	85.3	73.3	86.1	93.3	62.1	59.2	61.1	72.4
	F_1	97.3	70.7	53.6	48.7	69.4	65.8	95.4	65.4	57.4	60.4	86.2	85.2	72.3	85.9	93.7	59.8	57.5	63.4	71.6
M_6	REC	94.5	69.9	59.2	51.0	58.9	67.2	94.4	73.5	48.7	57.8	76.5	88.3	64.5	80.4	88.8	62.7	51.0	63.8	69.5
	PRE	97.8	70.6	59.6	52.3	63.4	70.3	93.4	72.4	46.4	62.2	82.2	88.8	66.9	83.4	93.4	67.5	57.6	65.2	71.9
	F_1	96.1	70.2	59.4	51.6	61.0	68.7	93.9	72.8	47.3	59.9	79.2	88.6	65.6	81.9	91.0	65.0	54.0	64.5	70.6
M_7	REC	97.6	72.2	47.1	54.5	71.6	73.0	93.3	45.8	52.9	61.4	77.4	78.5	70.0	81.9	97.4	75.9	66.3	59.6	70.9
	PRE	97.1	71.4	47.6	59.2	74.0	74.4	94.2	50.5	57.9	66.5	80.9	79.1	68.2	84.7	96.2	69.6	63.0	62.4	72.0
	F_1	97.3	71.8	47.3	56.7	72.8	73.6	93.8	47.9	55.3	63.7	79.0	78.7	69.0	83.2	96.8	72.5	64.6	60.9	71.4
M_8	REC	94.3	65.8	58.6	52.4	66.5	79.4	92.3	62.4	61.3	64.5	77.9	76.1	75.1	79.3	99.3	55.0	61.3	53.7	70.8
	PRE	96.6	68.2	58.8	53.3	65.0	78.6	94.6	63.7	62.0	67.6	80.8	79.0	76.8	79.7	98.2	60.4	60.8	56.4	72.2
	F_1	95.4	66.9	58.6	52.8	65.7	78.9	93.4	63.0	61.6	65.9	79.3	77.5	75.9	79.4	98.8	57.6	60.9	54.9	71.5
M_9	REC	97.7	69.8	54.6	43.4	66.3	76.2	91.8	55.8	52.9	78.9	92.7	87.2	71.7	81.9	98.1	71.6	49.3	53.3	71.8
	PRE	97.9	75.5	57.4	48.8	67.2	79.4	94.5	60.1	53.2	74.5	91.1	89.1	75.7	82.4	97.8	72.0	52.9	53.7	73.5
	F_1	97.8	72.5	55.9	45.9	66.7	77.6	93.1	57.8	53.0	76.6	91.9	88.1	73.7	82.1	97.9	71.8	51.0	53.5	72.6
M_{10}	REC	97.0	69.7	45.2	63.0	61.5	68.3	93.5	74.6	46.5	62.1	86.0	74.9	62.4	72.4	96.3	56.2	56.0	43.9	68.3
	PRE	97.0	71.8	44.5	60.9	64.7	72.6	95.3	71.5	50.0	65.3	86.0	78.6	66.0	73.0	97.5	55.5	57.0	46.3	69.6
	F_1	97.0	70.8	44.8	61.9	63.1	70.2	94.4	73.0	48.1	63.6	86.0	76.7	64.1	72.6	96.9	55.8	56.5	45.0	68.9
M_{11}	REC	97.7	59.0	60.2	56.5	72.3	79.5	91.5	55.4	56.5	62.8	79.0	76.0	64.6	72.7	92.6	58.8	44.5	47.8	68.2
	PRE	98.6	65.1	61.5	58.1	76.0	78.5	89.1	57.0	55.3	64.4	83.1	78.2	67.5	80.8	95.1	62.6	48.1	50.5	70.5
	F_1	98.1	61.9	60.8	57.2	74.0	79.0	90.2	56.2	55.7	63.6	80.9	77.0	66.0	76.5	93.8	60.6	46.2	49.1	69.3
Average across classes	REC	97.3	70.4	56.2	58.2	70.8	72.5	93.3	63.2	59.4	67.2	83.0	84.9	73.0	82.7	95.4	66.3	59.0	58.1	72.8
	PRE	97.4	71.0	56.6	58.5	71.4	73.1	93.5	63.3	59.5	67.7	83.3	85.1	73.4	83.2	95.5	66.8	59.2	58.3	73.2
	F_1	97.3	70.6	56.4	58.3	71.0	72.7	93.4	63.2	59.3	67.3	83.0	84.9	73.1	82.9	95.4	66.4	59.0	58.1	72.9

TABLE 4. The results obtained for each transradial amputated subject expressed in terms of the values of the REC, PRE, and F_1 evaluation metrics computed for each MI task.

| Subject	Metric	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}	Average across classes
A_1	REC	99.46	94.56	89.01	90.12	81.18	87.50	88.02	84.97	78.74	78.85	97.33	88.16
	PRE	94.73	93.32	90.31	88.58	83.09	90.27	89.77	88.98	78.10	79.56	93.24	88.18
	F_1	97.04	93.92	89.64	89.34	82.09	88.84	88.83	86.92	78.39	79.16	95.23	88.13
A_2	REC	77.43	79.76	79.76	72.13	75.40	70.99	76.36	75.68	71.33	63.84	79.50	74.74
	PRE	73.59	75.53	76.96	70.26	80.59	69.25	75.23	77.56	78.80	68.81	81.70	75.30
	F_1	75.45	77.56	78.32	71.18	77.87	70.04	75.78	76.58	74.84	66.08	80.54	74.93
A_3	REC	90.41	75.51	78.17	81.57	71.06	47.84	64.68	59.49	49.19	56.52	55.38	66.35
	PRE	80.64	68.85	72.95	79.24	65.63	55.04	65.74	62.60	59.14	63.35	58.05	66.47
	F_1	85.21	72.00	75.42	80.38	68.21	51.15	65.16	61.00	53.67	59.63	56.64	66.22
A_4	REC	60.00	66.72	65.24	57.02	56.65	60.46	56.32	55.45	65.54	51.36	56.49	59.20
	PRE	61.05	64.06	61.86	55.17	57.48	62.93	57.51	60.80	65.61	53.40	55.70	59.60
	F_1	60.48	65.35	63.50	56.00	57.02	61.59	56.87	58.00	65.56	52.29	56.07	59.34
Average across subjects	REC	81.82	79.14	78.04	75.21	71.07	66.70	71.35	68.90	66.20	62.64	72.17	72.11
	PRE	77.50	75.44	75.52	73.31	71.70	69.37	72.06	72.48	70.41	66.28	72.17	72.39
	F_1	79.55	77.21	76.72	74.22	71.30	67.91	71.66	70.63	68.11	64.29	72.12	72.16

amputated subject. The average $\pm$ STD CA value computed over all MI tasks and transradial amputated subjects is 72.8% $\pm$ 0.6%.

Table 4 shows the values of the REC, PRE, and F_1 metrics computed for each MI task per each transradial amputated subject. Moreover, the last column in Table 4, provides the average values of the REC, PRE, and F_1 metrics computed over all MI tasks per each transradial amputated subject. The last four rows in Table 4, which are associated with the columns labeled as M_1 to M_{11}, provides the average values of the REC, PRE, and F_1 metrics computed over all able-bodied subjects for each MI task. Finally, the last four rows in the last column of Table 4 provide the average values of the REC,

PRE, and F_1 metrics computed over all MI tasks and across all transradial amputated subjects. In particular, the overall average values of the REC, PRE, and F_1 metrics computed over all MI tasks and all transradial amputated subjects are 72.1%, 72.4%, and 72.2%, respectively.

C. RUNTIME OF OUR PROPOSED FRAMEWORK

The runtime results of our proposed framework are summarized in Table 5. In particular, Table 5 shows that the average $\pm$ STD training time that is computed over all able-bodied subjects is 1232.9 $\pm$ 14.79 s. Moreover, the average $\pm$ STD training time computed for all transradial amputate
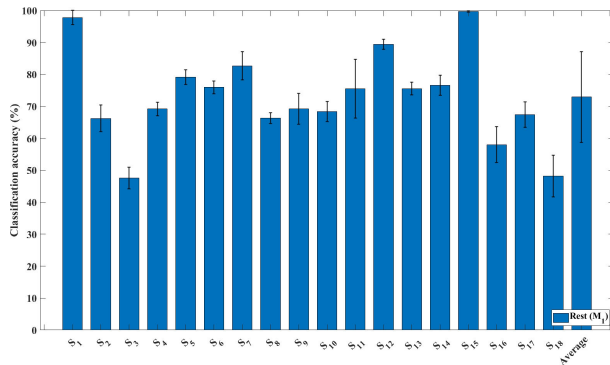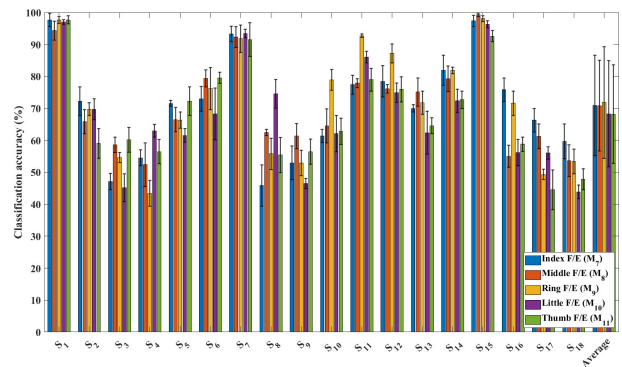
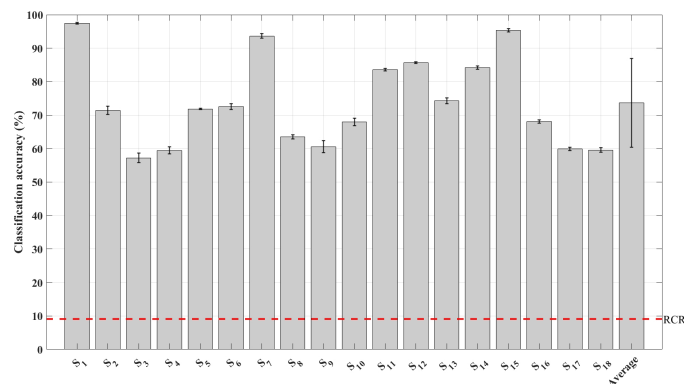(a) The CA values computed for the rest MI task per each able-bodied subject.
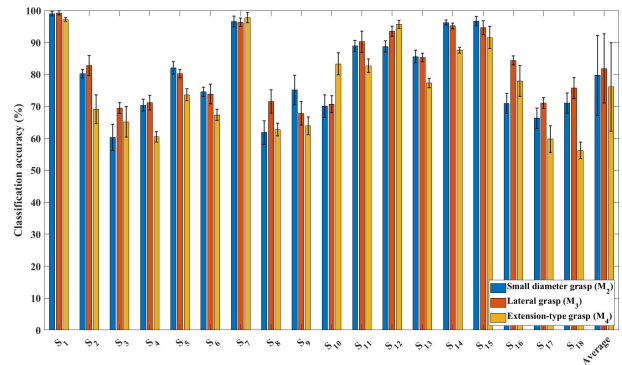
(b) The CA values computed for the grasp-related MI tasks per each able-bodied subject.
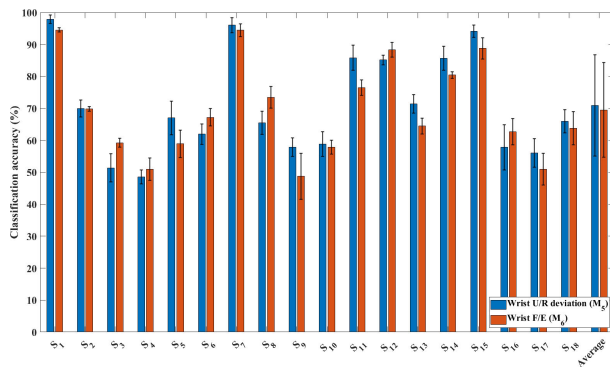
(c) The CA values computed for the wrist-related MI tasks per each able-bodied subject.

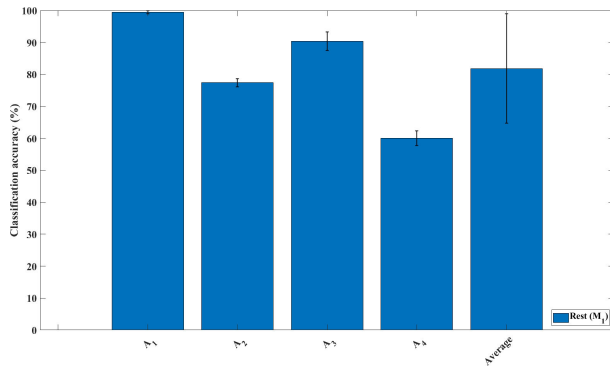(d) The CA values computed for the fingers-related MI tasks per each able-bodied subject.

(e) The CA values computed over all MI tasks per each able-bodied subject.

FIGURE 8. The results obtained using the EEG signals of the able-bodied subjects expressed in terms of the CA metric. The black vertical bars represent the STD values and the red dashed line represents the random classification rate (RCR), which is defined as the reciprocal of the number of MI tasks and equal to 9.09%.

subject is 1300.1 ± 10.7 s. The average $\pm$ STD time required to construct the image TFI_i at the input stage computed for all MI tasks over all able-bodied subjects is 0.0419 ± 0.0015 s, while the average $\pm$ STD time required to construct the image TFI_i for all MI tasks over all transradial amputated subjects is 0.0419 ± 0.0018 s. Finally, the average $\pm$ STD classification time computed for all MI tasks over all able-bodied subjects is 0.00059 ± 0.00001 s. In addition, the average $\pm$ STD classification time computed for all MI tasks over all transradial amputated subjects is 0.00054 ± 0.00003 s.

D. COMPARISON WITH THE CLASSIFICATION PERFORMANCE OBTAINED USING CONVENTIONAL CWD-BASED HANDCRAFTED FEATURES AND A MULTI-CLASS SVM CLASSIFIER

As described in section II, the feature extraction stage of our proposed framework learns latent features from the constructed TFI at each window position, namely TFI_i. These learned features are used at the classification stage to classify each input TFI into one of the eleven MI tasks. In this section, we compare the classification performance of
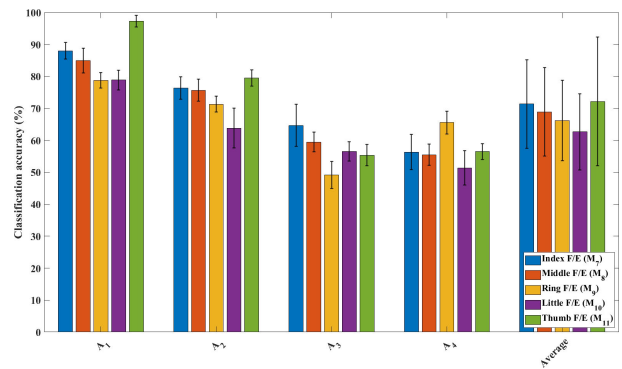
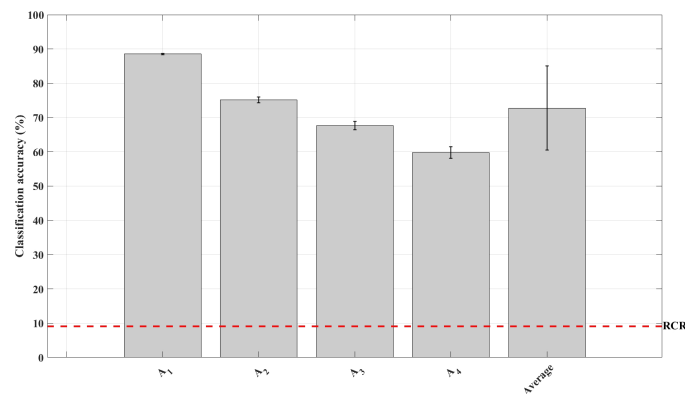(a) The CA values computed for the rest MI task per each transradial amputated subject.
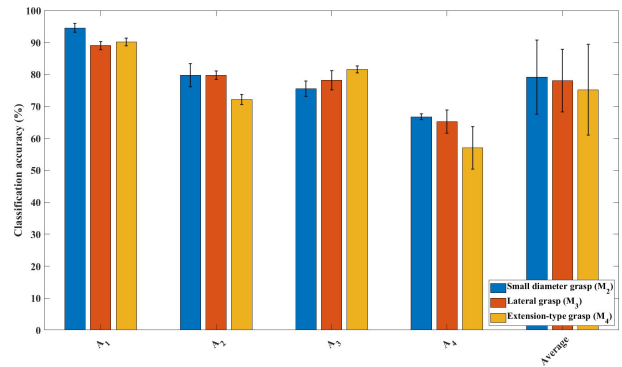
(b) The CA values computed for the grasp-related MI tasks per each transradial amputated subject.
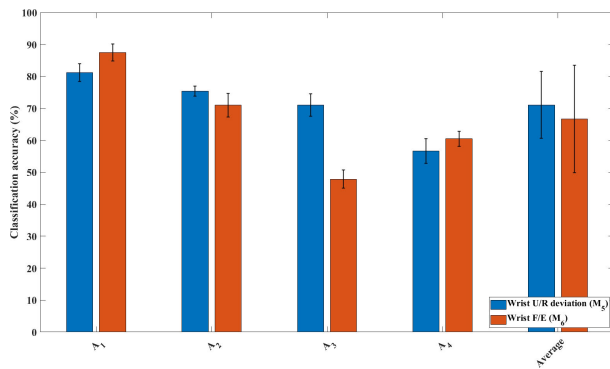
(c) The CA values computed for the wrist-related MI tasks per each transradial amputated subject.

(d) The CA values computed for the fingers-related MI tasks per each transradial amputated subject.

(e) The CA values computed over all MI tasks per each transradial amputated subject.
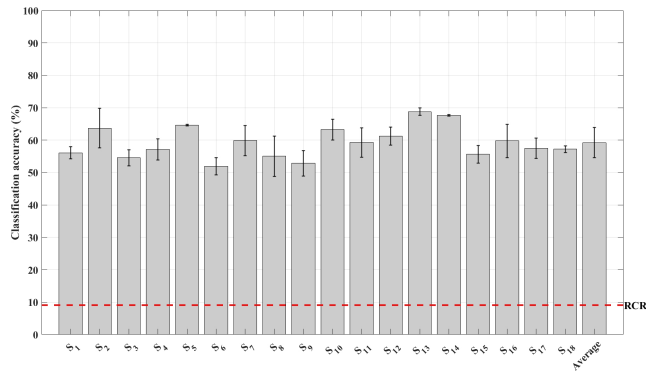
FIGURE 9. The results obtained using the EEG signals of the transradial amputated subjects expressed in terms of the CA metric. The black vertical bars represent the STD values and the red dashed line represents the random classification rate (RCR), which is defined as the reciprocal of the number of MI tasks and equal to 9.09%.

our proposed framework with the classification performance obtained using twelve conventional handcrafted features that are extracted from the CWD-based representation of the EEG signal [7]. In particular, the twelve conventional CWD-based handcrafted features, which are presented in [7], are re-implemented and utilized to construct a multi-class SVM classifier with radial basis function kernel to decode the eleven MI tasks. The performance of the constructed multi-class SVM classifier was evaluated using the ten-fold cross-validation procedure described in subsection II-D4.

Figures 10a and 10b show the average CA values obtained using the multi-class SVM classifier for the able-bodied and transradial amputated subjects, respectively. The average CA values presented in Fig. 10 are computed over the ten repetitions of the cross-validation procedure across the eleven MI tasks per each subject. In particular, the average $\pm$ STD CA value computed over all able-bodied subjects is 59.2% $\pm$ 4.6%, while the average $\pm$ STD CA value computed over all transradial amputated subjects is 61.6% $\pm$ 7.5%. The results presented in Fig. 10 indicate that the classification

TABLE 5. Runtime of our proposed framework.

Subjects	Average training time (s)	Average time to construct the image TFI_i (s)	Average time to classify the image TFI_i (s)
Able-bodied subjects	1232.9 ± 14.79	0.0419 ± 0.0015	0.00059 ± 0.00001
Transradial amputated subjects	1300.1 ± 10.7	0.0419 ± 0.0018	0.00054 ± 0.00003

(a) The CA values computed over all MI tasks per each able-bodied subject.
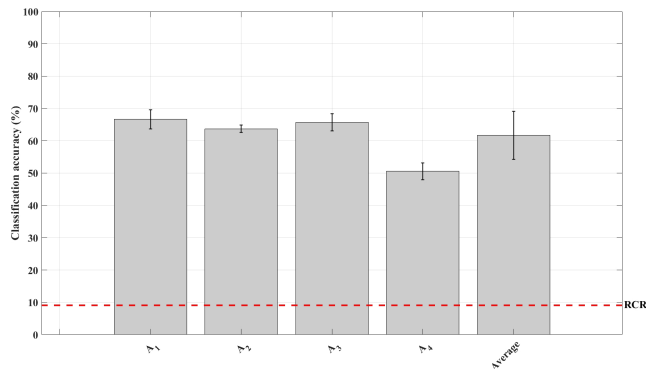
(b) The CA values computed over all MI tasks per each transradial amputated subject.

FIGURE 10. The CA values obtained using the conventional CWD-based handcrafted features and multi-class SVM classifier computed for the able-bodied and transradial amputated subjects. The black vertical bars represent the STD values and the red dashed line represents the RCR, which is equal to 9.09%.

performance of our proposed framework, which is shown in Fig. 8e and Fig. 9e, outperforms significantly the classification performance obtained using the CWD-based handcrafted features and a multi-class SVM classifier.

IV. DISCUSSION

In this work, we demonstrated the potential of utilizing a CNN to decode MI tasks within the same hand using the CWD-based TFIs that are extracted from the EEG signals. The performance of our proposed framework was evaluated using EEG data that were recorded for able-bodied and transradial amputated subjects while performing eleven MI tasks of the same hand. The results obtained for the able-bodied and transradial amputated subjects demonstrate the capability of our proposed framework to successfully decode eleven MI tasks of the same hand.

A. CLASSIFICATION PERFORMANCE OF OUR PROPOSED FRAMEWORK

Fig. 8e shows that the average CA value of each able-bodied subject, which is computed over all MI tasks, is considerably higher than the RCR, which is defined as the reciprocal of the number of MI tasks and its value is equal to 9.09%. Similarly, the results presented in Fig. 9e indicate that the average CA value computed for each transradial amputated subject over all MI tasks is considerably higher than the RCR. This implies that the input stage provides a suitable representation of the EEG signals in the joint time-frequency domain. Moreover, these results suggest the capability of the feature extraction stage to extract salient features from the TFIs, which are constructed at the input stage, that can be used to discriminate between different MI tasks of the same hand.

The results obtained for the able-bodied and transradial amputated subjects, which are presented in subsection III-A and subsection III-B, show substantial variation in the classification performance computed for different subjects. In particular, for the able-bodied subjects, the mean CA values computed for each subject over all MI tasks are between 59.5%, which was obtained for the subject S_4, and 97.4%, which was obtained for the subject S_1. Similarly, for the transradial amputated subjects, the mean CA values computed for each subject over the MI tasks are between 59.8%, which was obtained for the subject A_4, and 88.5%, which was obtained for the subject A_1. This variation in the classification performance obtained for the able-bodied and transradial amputated subjects can be attributed to the different capability of the subjects to successfully perform various MI tasks [6], [50].

The classification results obtained for the able-bodied and transradial amputated subjects, which are presented in Table 3 and Table 4, indicate that the REC, PRE, and F_1 values computed for the grasp-related MI tasks, namely M_2, M_3, and M_4, are relatively higher than the REC, PRE, and F_1 values computed for the wrist- and fingers-related MI tasks. This can be attributed to the following factors [22], [23]: (1) Grasp-related tasks activate larger brain regions compared to the activated brain regions during wrist- or fingers-related MI tasks. (2) Wrist- and fingers-related MI tasks activate relatively small and close regions in the sensorimotor cortex region. (3) The low SNR and limited spatial resolution of the EEG signals reduce the capability of distinguishing between brain activities recorded during wrist- and fingers-related MI tasks.

B. RUNTIME ANALYSIS

The runtime analysis presented in Table 5 indicates the possibility of utilizing our proposed framework to construct a real-time system for decoding MI tasks of the same hand. In particular, for able-bodied subjects, the average time required to construct and classify the image TFI_i using the trained framework is equal to 0.0478 s. We refer to the average time required to construct and classify the image TFI_i as the average response time. The ratio between the average response time and the duration of the sliding window, which is equal

to 0.250 s, is approximately 19.1%. Similarly, for transradial amputed subjects, the average response time computed using the trained framework is equal to 0.0424 s. Hence, the ratio between the average response time and the duration of the sliding window, which is equal to 0.250 s, is approximately 16.9%. These results indicate that our proposed framework can classify the EEG segments associated with each window position into one of the eleven MI tasks before the extraction of the next set of segments from the EEG signals (i.e., before moving from the current window position to the next window position). This implies the possibility of utilizing our proposed framework to construct a system that can discriminate between MI tasks of the same hand in real time.

V. CONCLUSION

In this paper, we proposed a novel framework that employs a new CNN architecture to decode MI tasks associated with fine body-parts within the same hand, such as the wrist and fingers, using a CWD-based representation of the EEG signals. The performance of our proposed framework was evaluated using EEG signals that were recorded from eighteen able-bodied subjects and four transradial amputed subjects while performing eleven MI tasks within the same hand. The results presented in this study indicate the capability of our proposed framework to distinguish between the MI tasks within the same hand for able-bodied and transradial amputed subjects.

In the future, we plan to investigate the possibility of applying our proposed deep learning frame to decode movements and imaginary tasks that are performed by the same upper extremity, such as the elbow and shoulder. In addition, we intend to compare the performance of our proposed deep learning framework with the performance obtained using other traditional state-of-the-art feature extraction procedures, such as the common spatial patterns.

REFERENCES

[1] H. A. Agashe, A. Y. Paek, and J. L. Contreras-Vidal, "Multisession, noninvasive closed-loop neuroprosthetic control of grasping by upper limb amputees," *Prog. Brain Res.*, vol. 228, p. 107, Jan. 2016.

[2] L. Resnik, M. R. Meucci, S. Lieberman-Klinger, C. Fantini, D. L. Kelty, R. Disla, and N. Sasson, "Advanced upper limb prosthetic devices: Implications for upper limb prosthetic rehabilitation," *Arch. Phys. Med. Rehabil.*, vol. 93, no. 4, pp. 710–717, 2012.

[3] Z. Tang, S. Sun, S. Zhang, Y. Chen, C. Li, and S. Chen, "A brain-machine interface based on ERD/ERS for an upper-limb exoskeleton control," *Sensors*, vol. 16, no. 12, p. 2050, 2016.

[4] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[5] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[6] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, 2017, Art. no. 016003.

[7] R. Alazrai, H. Alwanni, Y. Baslan, N. Alnuman, and M. I. Daoud, "Eeg-based brain–computer interface for decoding motor imagery tasks within the same hand using Choi–Williams time-frequency distribution," *Sensors*, vol. 17, no. 9, p. 1937, 2017.

[8] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper, "Brain oscillations control hand orthosis in a tetraplegic," *Neurosci. Lett.*, vol. 292, no. 3, pp. 211–214, 2000.

[9] M. Jeannerod, "Mental imagery in the motor context," *Neuropsychologia*, vol. 33, no. 11, pp. 1419–1432, 1995.

[10] R. Gatti, A. Tettamanti, P. M. Gough, E. Riboldi, L. Marinoni, and G. Buccino, "Action observation versus motor imagery in learning a complex motor task: A short review of literature and a kinematics study," *Neurosci. Lett.*, vol. 540, pp. 37–42, Apr. 2013.

[11] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain–computer communication," *Proc. IEEE*, vol. 89, no. 7, pp. 1123–1134, Jul. 2001.

[12] 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, 2018, Art. no. 056013.

[13] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for brain mapping and decoding of movement-related information from the human EEG," 2017, *arXiv:1703.05051v1*. [Online]. Available: https://arxiv.org/abs/1703.05051v1

[14] H. Dose, J. S. Møller, H. K. Iversen, and S. Puthusserypady, "An end-to-end deep learning approach to MI-EEG signal classification for BCIs," *Expert Syst. Appl.*, vol. 114, pp. 532–542, Dec. 2018.

[15] A. J. Doud, J. P. Lucas, M. T. Pisansky, and B. He, "Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain–computer interface," *PLoS ONE*, vol. 6, no. 10, 2011, Art. no. e26322.

[16] A. S. Royer, A. J. Doud, M. L. Rose, and B. He, "EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 6, pp. 581–589, Dec. 2010.

[17] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface," *J. Neural Eng.*, vol. 10, no. 4, 2013, Art. no. 046003.

[18] F. Quandt, C. Reichert, H. Hinrichs, H. J. Heinze, R. T. Knight, and J. W. Rieger, "Single trial discrimination of individual finger movements on one hand: A combined MEG and EEG study," *NeuroImage*, vol. 59, no. 4, pp. 3316–3324, 2012.

[19] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 4–14, Jan. 2016.

[20] A. Vuckovic and F. Sepulveda, "Delta band contribution in cue based single trial classification of real and imaginary wrist movements," *Med. Biol. Eng. Comput.*, vol. 46, no. 6, pp. 529–539, 2008.

[21] X. Yong and C. Menon, "EEG classification of different imaginary movements within the same limb," *PLoS ONE*, vol. 10, no. 4, Apr. 2015, Art. no. e0121896.

[22] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, "Decoding individual finger movements from one hand using human EEG signals," *PLoS ONE*, vol. 9, no. 1, 2014, Art. no. e85192.

[23] R. Alazrai, H. Alwanni, and M. I. Daoud, "EEG-based BCI system for decoding finger movements within the same hand," *Neurosci. Lett.*, vol. 698, pp. 113–120, Apr. 2019.

[24] G. Pfurtscheller and F. L. Da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, 1999.

[25] B. Boashash, *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. New York, NY, USA: Academic, 2015.

[26] 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.

[27] M. Grosse-Wentrup and M. Buss, "Multiclass common spatial patterns and information theoretic feature extraction," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1991–2000, Aug. 2008.

[28] R. Alazrai, M. Momani, H. A. Khudair, and M. I. Daoud, "EEG-based tonic cold pain recognition system using wavelet transform," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 3187–3200, Jul. 2019.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[30] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals," *IEEE Access*, vol. 6, pp. 25399–25410, 2018.

[31] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, May 2015.

[32] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[33] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, 2018.

[34] B. Blankertz, K.-R. Muller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlogl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroder, and N. Birbaumer, "The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1044–1051, Jun. 2004.

[35] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. R. Müller-Putz, G. Nolte, G. Pfurtscheller, H. Preissl, G. Schalk, A. Schlögl, C. Vidaurre, S. Waldert, and B. Blankertz, "Review of the BCI competition IV," *Frontiers Neurosci.*, vol. 6, no. 1, p. 55, 2012.

[36] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[37] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain–computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.

[38] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, S. Elsig, G. Giatsidis, F. Bassetto, and H. Müller, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, Dec. 2014, Art. no. 140053.

[39] S.-M. Zhou, J. Q. Gan, and F. Sepulveda, "Classifying mental tasks based on features of higher-order statistics from EEG signals in brain–computer interface," *Inf. Sci.*, vol. 178, no. 6, pp. 1629–1640, 2008.

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

[41] G. Gómez-Herrero, W. De Clercq, H. Anwar, O. Kara, K. Egiazarian, S. Van Huffel, and W. Van Paesschen, "Automatic removal of ocular artifacts in the EEG without an EOG reference channel," in *Proc. IEEE 7th Nordic Signal Process. Symp.*, Jun. 2006, pp. 130–133.

[42] P. Castiglioni, *Choi-Williams Distribution*. Hoboken, NJ, USA: Wiley, 2005.

[43] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, Jun. 1989.

[44] J. M. O'Toole, "Discrete quadratic time-frequency distributions: Definition, computation, and a newborn electroencephalogram application," Ph.D. dissertation, School Med., Univ. Queensland, Brisbane, QLD, Australia, 2009.

[45] R. Alazrai, R. Homoud, H. Alwanni, and I. M. Daoud, "EEG-based emotion recognition using quadratic time-frequency distribution," *Sensors*, vol. 18, no. 8, p. 2739, Aug. 2018.

[46] S. L. Hahn, *Hilbert Transforms in Signal Processing*, vol. 2. Boston, MA, USA: Artech House, 1996.

[47] L. Cohen, "Time-frequency distributions-a review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, Jul. 1989.

[48] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2017.

[49] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.

[50] M. Ahn and S. C. Jun, "Performance variation in motor imagery brain–computer interface: A brief review," *J. Neurosci. Methods*, vol. 243, pp. 103–110, Mar. 2015.

RAMI ALAZRAI received the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 2013. In June 2013, he joined the School of Electrical Engineering and Information Technology, German Jordanian University (GJU), where he is currently an Associate Professor. His research interests include biomedical signal processing, brain–computer interfaces, decoding actual and imaginary motor tasks using EEG signals, EEG-based emotion identification, EEG-based pain analysis, semantic video analysis, human activity and emotion analysis, human–human interaction representation and analysis, and elderly fall detection.

MOTAZ ABUHIJLEH received the B.S. degree in computer engineering from German Jordanian University (GJU), Amman, Jordan, in 2019. As an obligatory part of his bachelor's program, he spent a year at the HS Bremen City University of Applied Sciences, Germany. He is currently working as a Research Assistant on several research projects in the fields of biomedical signal processing, machine learning, and brain–computer interfaces.

HISHAM ALWANNI received the B.S. degree in mechatronics engineering from German Jordanian University, Amman, Jordan, in 2016. He is currently pursuing the master's degree with the Faculty of Engineering, University of Freiburg, Germany. His research interests include EEG signal processing, brain–computer interfaces, deep learning in medical image analysis, and virtual reality.

MOHAMMAD I. DAOUD received the Ph.D. degree in electrical and computer engineering from the University of Western Ontario, London, Canada, in 2009. He was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada, where he held an NSERC Postdoctoral Fellowship. In September 2011, he joined the Department of Computer Engineering, German Jordanian University, Amman, Jordan, where he is currently an Associate Professor. His research interests include medical image and signal processing, ultrasound imaging, machine learning, and parallel computing. During his graduate studies, he was awarded the NSERC PGS-D Scholarship and has served as a Scholar of the Canadian Institutes of Health Research/University of Western Ontario Strategic Training Initiative in Cancer Research and Technology Transfer.

• • •