Motor Imagery Classification for Asynchronous EEG-Based Brain–Computer Interfaces

Huanyu Wu, Siyang Li, and Dongrui Wu, *Fellow, IEEE*

Abstract— **Motor imagery (MI) based brain-computer interfaces (BCIs) enable the direct control of external devices through the imagined movements of various body parts. Unlike previous systems that used fixed-length EEG trials for MI decoding, asynchronous BCIs aim to detect the user's MI without explicit triggers. They are challenging to implement, because the algorithm needs to first distinguish between resting-states and MI trials, and then classify the MI trials into the correct task, all without any triggers. This paper proposes a sliding window prescreening and classification (SWPC) approach for MI-based asynchronous BCIs, which consists of two modules: a prescreening module to screen MI trials out of the resting-state, and a classification module for MI classification. Both modules are trained with supervised learning followed by self-supervised learning, which refines the feature extractors. Within-subject and cross-subject asynchronous MI classifications on four different EEG datasets validated the effectiveness of SWPC, i.e., it always achieved the highest average classification accuracy, and outperformed the best state-of-the-art baseline on each dataset by about 2%.**

Index Terms— **Brain-computer interface, electroencephalogram, motor imagery, self-supervised learning.**

I. INTRODUCTION

NON-INVASIVE electroencephalogram (EEG) based brain-computer interfaces (BCIs) have made rapid progress recently [1], [2], [3]. As shown in Fig. 1, a closed-loop BCI system usually consists of three components: signal acquisition, signal analysis, and external device control. Signal analysis further includes signal preprocessing, feature extraction, and classification. Both traditional classifiers, e.g., linear discriminant analysis and support vector machine, and deep learning, e.g., EEGNet [4], have been used.

There are three classical BCI paradigms: motor imagery (MI), event-related potential, and steady-state visual evoked
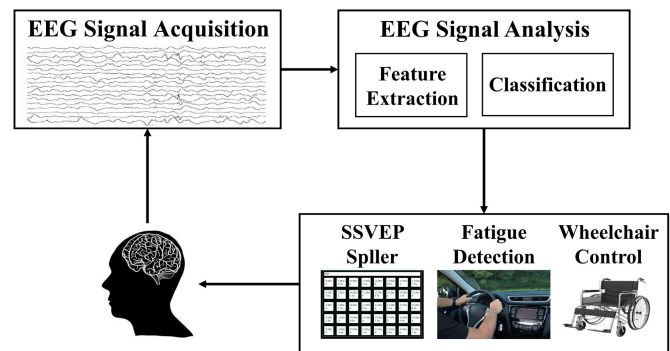
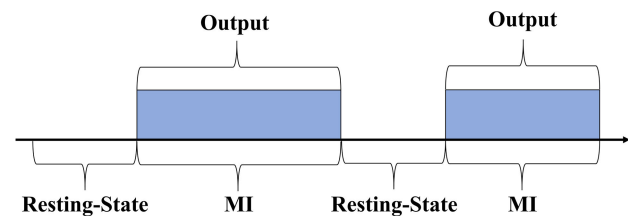Fig. 1. Flowchart of a closed-loop EEG-based BCI system.

Fig. 2. Illustration of asynchronous MI classification. The user may switch between resting-state and MI at any unknown time.

potential. The paper focuses on MI, where the user imagines the movement of various body parts, e.g., left/right hand, both feet or tongue, to elicit different EEG patterns and hence to control external devices. It has been used in upper limb robotic rehabilitation [5], text input [6], wheelchair control [7], etc.

Most existing MI-based BCIs use specific triggers to indicate the start and end of each MI trial, which may be inconvenient in practice. For example, when an MI-based BCI is used to navigate a wheelchair, the control commands should be sent out whenever the user wants, instead of only after some triggers.

Mason and Birch [8] first proposed the concept of asynchronous BCIs. As shown in Fig. 2, a subject generates control signals by consciously changing his/her mental state: when the subject starts MI, the BCI system detects it and completes the corresponding instruction; otherwise, it keeps still.

Asynchronous MI-based BCIs are challenging to implement, because the algorithm needs to first distinguish between resting-states (no MI) and MI trials, and then classify the MI trials into the correct task, all without any triggers. Very few studies have appeared in the literature. For example,
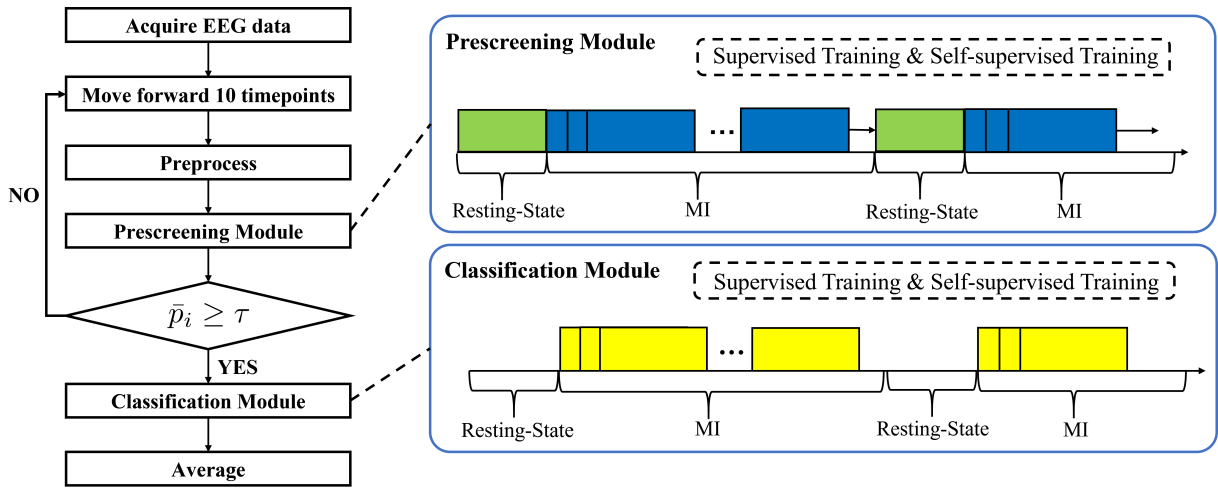
Fig. 3.   SWPC for asynchronous MI-based BCIs.

Sugiura et al. [9] adopted a hierarchical hidden Markov model, and Saa and Cetin [10] proposed conditional random fields and latent dynamic conditional random fields for EEG classification in asynchronous BCIs.

This paper proposes a sliding window prescreening and classification (SWPC) approach for asynchronous MI-based BCIs, which consists of two modules:

1) *Prescreening module*, where a classifier with a fixed window length, trained with both supervised learning and self-supervised learning (SSL), is used to prescreen MIs from the resting-state. If the output probability exceeds a threshold, then the EEG trial is sent to the next module for classification.

2) *Classification module*, where a classifier, also trained with supervised learning and SSL, is used for MI classification.

Within-subject and cross-subject experiments on four MI datasets demonstrated the effectiveness of SWPC, particularly SSL, to refine the feature extractors.

The remainder of this paper is organized as follows. Section II introduces our proposed SWPC. Section III validates the performance of SWPC on four MI datasets. Finally, Section IV draws conclusions and points out future research directions.

II. METHODOLOGY

This section introduces our proposed SWPC for asynchronous MI-based BCIs. The code is publicly available at https://github.com/why135724/SWPC.

A. Flowchart of SWPC

Fig. 3 shows the flowchart of our proposed SWPC for asynchronous MI-based BCIs. It includes two modules:

1) *Prescreening module*, where an EEGNet [4] classifier with a fixed window length is trained to screen the MIs out of the resting-state EEG trials.

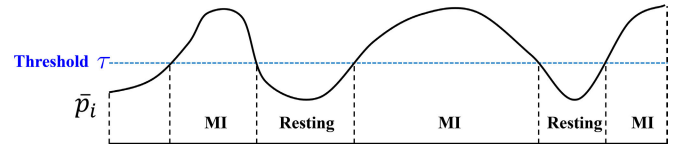2) *Classification module*, where another EEGNet classifier is trained to classify the prescreened MI trials.

Fig. 4.   Usage of the prescreening probability $\bar{p}_i$.

B. Problem Setting

Two training sets are used in SWPC.

The first training set, $\mathcal{D}^s = \{(X_i^s, y_i^s)\}_{i=1}^{n_s}$, is used in the classification module to classify potential MI trials into different MI tasks. It consists of n_s labeled MI trials $X_i^s \in \mathbb{R}^{ch \times ts}$ and the corresponding labels y_i^s, where ch is the number of EEG channels, and ts the number of time domain samples.

The second training set, $\bar{\mathcal{D}}^s = \{(\bar{X}_i^s, \bar{y}_i^s)\}_{i=1}^{2n_s}$, is used in the prescreening module to distinguish MI trials from the resting-state. It consists of n_s labeled MI trials from $\mathcal{D}^s$, and another n_s resting-state trials $\bar{X}_i^s \in \mathbb{R}^{ch \times ts}$ adjacent to the MI trials. $\bar{y}_i^s \in \{0, 1\}$ (0 denotes resting-state, and 1 denotes MI) is the label of $\bar{X}_i^s$.

The test data $X^t \in \mathbb{R}^{ch \times fl}$ is a long EEG data stream with fl time domain samples, where usually $fl \gg ts$. It does not include any triggers, so we do not know when an MI trial starts. The goal is to correctly identify the MI periods and further classify them into specific MI tasks.

To simplify the problem, we split X^t with sliding window length L_w and step 10 to get n_t test trials $\mathcal{D}_t = \{X_i^t\}_{i=1}^{n_t}$. Each trial X_i^t is then passed to the prescreening module, which outputs the probability $\bar{p}_i$ of X_i^t being MI. If $\bar{p}_i$ exceeds a threshold τ, as shown in Fig. 4, then X_i^t is further passed to the classification module. If multiple successive trials have $\bar{p}_i \geq \tau$, then their corresponding classification probabilities are averaged as the final output. The predicted label for X_i is denoted as $\hat{y}_i^t$.

C. The Prescreening Module

As shown in Fig. 5, the prescreening module includes first supervised training and then SSL.
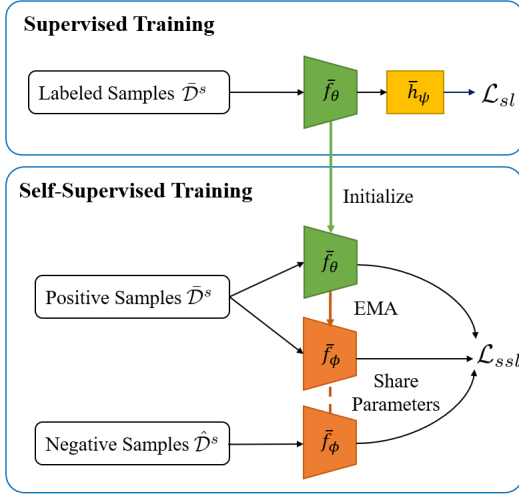
Fig. 5.   Training of the prescreening module.

Fig. 6.   Training of the classification module.

Supervised training performs binary classification between MI and resting-states. It trains a feature extractor $\bar{f}_{\theta}$ and a classifier $\bar{h}_{\psi}$ on $\bar{\mathcal{D}}^s$, using the cross-entropy loss:

$$\mathcal{L}_{sl}(\bar{h}_{\psi}, \bar{f}_{\theta}) = -\sum_{i=1}^{2n_s} \bar{y}_i^s \log\left(\bar{h}_{\psi}(\bar{f}_{\theta}(\bar{X}_i^s))\right). \tag{1}$$

Both $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ are updated with gradient descent.

SSL is used to fine-tune the feature extractor $\bar{f}_{\theta}$. It first constructs $2n_s$ transition trials (negative samples):

$$\hat{X}_i^s = 0.5(\bar{X}_r^s + \bar{X}_m^s), \quad i = 1, \ldots, 2n_s \tag{2}$$

where $\bar{X}_r^s$ and $\bar{X}_m^s$ are randomly selected resting-state trial and MI trial from $\bar{\mathcal{D}}^s$, respectively. Next, it updates the feature extractor $\bar{f}_{\theta}$ and simultaneously another auxiliary feature extractor $\bar{f}_{\phi}$, on positive samples $\bar{\mathcal{D}}^s$ and negative samples $\hat{\mathcal{D}}^s = \{\hat{X}_i\}_{i=1}^{n_s}$ using the contrastive loss:

$$\mathcal{L}_{ssl}(\bar{f}_{\theta}, \bar{f}_{\phi}) = \delta \cdot \exp\left(-\frac{\sum_{i=1}^{2n_s} |\bar{f}_{\theta}(\hat{X}_i^s) - \bar{f}_{\phi}(\hat{X}_i^s)|^2}{2\sigma^2}\right)$$
$$- \exp\left(-\frac{\sum_{i=1}^{2n_s} |\bar{f}_{\theta}(\bar{X}_i^s) - \bar{f}_{\phi}(\bar{X}_i^s)|^2}{2\sigma^2}\right), \tag{3}$$

where $\delta = 0.3$ is a hyperparameter controlling the contribution of the negative samples, and $\sigma = 2.0$ determines the Gaussian kernel width. Note that $\bar{f}_{\theta}$ and $\bar{f}_{\phi}$ are L2-normalized before entering (3), i.e.,

$$\bar{f}_{\theta}(\bar{X}_i^s) \leftarrow \frac{\bar{f}_{\theta}(\bar{X}_i^s)}{\|\bar{f}_{\theta}(\bar{X}_i^s)\|_2}, \quad \bar{f}_{\phi}(\hat{X}_i^s) \leftarrow \frac{\bar{f}_{\phi}(\bar{X}_i^s)}{\|\bar{f}_{\phi}(\bar{X}_i^s)\|_2}. \tag{4}$$

$\bar{f}_{\theta}$ in (3) is optimized by gradient descent, whereas $\bar{f}_{\phi}$ is optimized through exponential moving average (EMA):

$$\boldsymbol{\phi}^{n+1} \leftarrow \lambda \cdot \boldsymbol{\phi}^n + (1 - \lambda) \cdot \boldsymbol{\theta}^n, \tag{5}$$

where $\lambda = 0.9995$.
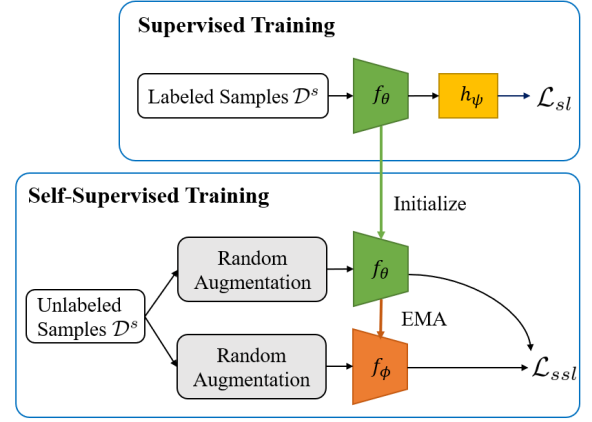
For a test EEG trial X_i^t, only $\bar{f}_{\theta}$ and $\bar{h}_{\psi}$ are used to compute the prescreening probability $\bar{p}_i$, i.e.,

$$\bar{p}_i = \bar{h}_{\psi}(\bar{f}_{\theta}(X_i^t)). \tag{6}$$

If $\bar{p}_i$ exceeds a threshold τ, then the EEG trial is further passed to the classification module.

D. The Classification Module

As shown in Fig. 6, the training process of the classification module is similar to that of the prescreening module. It also consists of two steps: supervised training and SSL.

Supervised training of the classification module remains the same as supervised training of the prescreening module, except that the EEG trials are classified into different MI tasks, instead of MI and resting-state.

SSL is again used to refine the feature extractor f_{θ}. We replace the construction of negative samples in the prescreening module with data augmentation in the classification module. The following data augmentations are used in this paper:
1) Adding noise: Uniform noise $0.5U(-\delta, \delta)$ is added to each element of the feature vector, where δ is the standard deviation of the original feature.
2) Scaling amplitude: Each feature is scaled by 0.75 or 1.25.
3) Masking channels: Randomly set all signals in some EEG channels to 0.
4) Masking segments: Randomly set some segments of the EEG signal to 0.

For each trial X_i^s, we randomly select two different data augmentations to get $X_{i,1}^s$ and $X_{i,2}^s$. $\{X_{i,1}^s\}_{i=1}^{n_s}$ and $\{X_{i,2}^s\}_{i=1}^{n_s}$ are then L2-normalized using (4) before computing the following contrastive loss:

$$\mathcal{L}_{ssl}(f_{\theta}, f_{\phi}) = -\exp\left(-\frac{\sum_{i=1}^{n_s} |f_{\theta}(X_{i,1}^s) - f_{\phi}(X_{i,2}^s)|^2}{2\sigma^2}\right). \tag{7}$$

For an input test trial X_i^t, the instantaneous classification probability is

$$p_i = h_{\psi}(f_{\theta}(X_i^t)). \tag{8}$$

Fig. 7.   Illustration of computing $\hat{p}_i$.

TABLE I
SUMMARY OF THE FOUR MI DATASETS

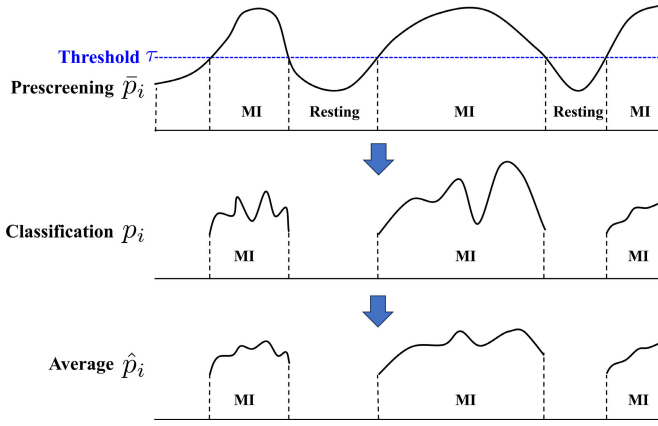| Dataset	Number of Subjects	Number of Channels	Trials per Subject	Number of Classes
MI1	9	22	144	2
MI2	9	22	288	4
MI3	14	15	100	2
MI4	9	3	60	2

To stabilize the output, we average all successive p_i whose corresponding $\bar{p}_i$ exceed τ, i.e.,

$$\hat{p}_i = \frac{1}{i - i_0} \sum_{j=i_0}^{i} p_j, \qquad (9)$$

where i_0 is the smallest index that ensures all $\{\bar{p}_j\}_{j=i_0}^{i}$ exceed the threshold τ. $\hat{p}_i$ is the final prediction probability for X_i^t.

More specifically, as shown in Fig. 7, the process of computing $\hat{p}_i$ is: We pass $X_i^t \in \mathcal{D}_t$ to $\bar{f}_\theta$ and $\bar{h}_\psi$ to get $\bar{p}_i$. If $\bar{p}_i < \tau$, then we classify the corresponding X_i^t as resting-state; otherwise, we further pass X_i^t to f_θ and h_ψ to get p_i. $\hat{p}_i$ is then averaged by (9) and used to derive $\hat{y}_i^t$.

The pseudo-code of SWPC is given in Algorithm 1.

III. Experiments

This section evaluates the performance of our proposed SWPC on four public MI datasets in both within-subject and cross-subject classification.

A. Datasets

Four public datasets from BNCI-Horizon,[1] summarized in Table I, were used in our experiments:

1) MI1 was the 001-2014 dataset [11] recorded from 9 subjects. Each session included 6 runs separated by short breaks. EEG signals were sampled at 250Hz. Only two classes (left-hand and right-hand) were used.
2) MI2 was also the 001-2014 dataset, but with all four classes, i.e., left-hand, right-hand, feet, and tongue.
3) MI3 was the 002-2014 dataset [12] recorded from 14 subjects. Each session included 8 runs separated by short breaks. EEG signals were sampled at 512Hz.

[1]http://www.bnci-horizon-2020.eu/database/data-sets

Algorithm 1 SWPC for Asynchronous MI-Based BCIs

Require: Training set $\mathcal{D}^s = \{(X_i^s, y_i^s)\}_{i=1}^{n_s}$;
    Training set $\bar{\mathcal{D}}_s = \{(\bar{X}_i^s, \bar{y}_i^s)\}_{i=1}^{2n_s}$;
    Maximum number of epochs, $\max_1$ and $\max_2$;
Ensure: The classification $\{\hat{y}_i^t\}_{i=1}^{n_t}$.
1: // Supervised Training of the Prescreening Module
2: Randomly initialize the feature extractor $\bar{f}_\theta$ and the classifier $\bar{h}_\psi$;
3: **for** $k = 1, \ldots, \max_1$ **do**
4:     Pass $\bar{\mathcal{D}}_s$ to $\bar{f}_\theta$ and $\bar{h}_\psi$ to compute $\mathcal{L}_{sl}(\bar{h}_\psi, \bar{f}_\theta)$ in (1);
5:     Update θ and ψ;
6: **end for**
7: // Self-supervised Training of the Prescreening Module
8: Initialize ϕ to θ;
9: Generate $\hat{X}^s$ using (2);
10: **for** $k = 1, \ldots, \max_2$ **do**
11:     Pass $\{X_i^s\}_{j=1}^{2n_s}$ and $\{\hat{X}_i^s\}_{i=1}^{2n_s}$ to $\bar{f}_\theta$ and $\bar{f}_\phi$ to compute $\mathcal{L}_{ssl}(\bar{f}_\theta, \bar{f}_\phi)$ in (3);
12:     Update θ, and ϕ using (5);
13: **end for**
14: // Supervised Training of the Classification Module
15: Randomly initialize the feature extractor f_θ and the classifier h_ψ;
16: **for** $k = 1, \ldots, \max_1$ **do**
17:     Pass $\mathcal{D}^s$ to f_θ and h_ψ to compute $\mathcal{L}_{sl}(h_\psi, f_\theta)$ in (1);
18:     Update θ and ψ;
19: **end for**
20: // Self-supervised Training of the Classification Module
21: Initialize ϕ to θ;
22: **for** $k = 1, \ldots, \max_2$ **do**
23:     Augment $\{X_i^s\}_{i=1}^{n_s}$ twice to get $\{X_{i,1}^s\}_{i=1}^{n_s}$ and $\{X_{i,2}^s\}_{i=1}^{n_s}$;
24:     Pass $\{X_{i,1}^s\}_{i=1}^{n_s}$ to f_θ and $\{X_{i,2}^s\}_{i=1}^{n_s}$ to f_ϕ, and compute $\mathcal{L}_{ssl}(f_\theta, f_\phi)$ in (7);
25:     Update θ, and ϕ using (5);
26: **end for**
27: // Test Sample Classification
28: Split X^t with sliding window length L_w and step size 10 to get $\mathcal{D}_t = \{X_i^t\}_{i=1}^{n_t}$;
29: Set $i_0 = 0$;
30: **for** $i = 1, \ldots, n_t$ **do**
31:     Pass X_i^t to $\bar{f}_\theta$ and $\bar{h}_\psi$ to get $\bar{p}_i$;
32:     **if** $\bar{p}_i \geq \tau$ **then**
33:         **if** $i_0 == 0$ **then**
34:             Set $i_0 = i$;
35:         **end if**
36:         Pass X_i^t to f_θ and h_ψ to compute p_i;
37:         Compute $\hat{p}_i$ using (9);
38:         Obtain $\hat{y}_i^t$ from $\hat{p}_i$;
39:     **else**
40:         Set $\hat{y}_i^t$ as resting-state;
41:         Set $i_0 = 0$;
42:     **end if**
43: **end for**

TABLE II
ACCs ON MI1 IN WITHIN-SUBJECT CLASSIFICATION

Approach	Subject							Average
	1	2	3	4	5	6	7	
CEC	61.35	51.57	82.76	52.75	45.24	85.45	77.64	65.25±16.47
JTS	67.34	51.88	87.34	55.62	50.24	90.01	82.57	69.29±17.26
BYOL	71.31	52.56	85.19	53.94	**54.64**	88.67	77.56	69.12±15.44
SimCLR	67.06	52.47	85.81	55.94	53.17	88.58	85.11	69.73±16.43
MoCo	65.28	54.17	**87.51**	56.25	54.17	89.58	**86.81**	70.54±16.75
ContraWR	68.45	53.49	85.79	56.78	51.76	87.65	84.65	69.80±16.11
SSCL	65.33	51.43	83.88	53.45	52.21	84.67	81.17	67.45±15.51
Ou2022	64.31	52.34	84.79	54.67	51.56	86.65	83.45	68.25±16.20
MAML	70.56	50.59	81.54	53.91	50.42	86.71	81.89	67.95±16.04
EOA	71.21	52.81	83.81	57.91	51.09	**91.41**	79.61	69.69±16.03
Song2022	71.57	**57.27**	85.31	55.12	51.89	86.51	85.21	70.41±15.55
SWPC	**73.61**	56.25	**87.51**	**58.33**	52.78	90.97	84.72	**72.02**±16.17

TABLE III
ACCs ON MI2 IN WITHIN-SUBJECT CLASSIFICATION

Approach	Subject							Average
	1	2	3	4	5	6	7	
CEC	37.12	25.45	48.23	31.45	35.45	52.14	50.24	40.01±10.28
JTS	42.37	27.45	51.78	37.69	38.31	54.25	51.27	43.30±9.69
BYOL	51.12	28.21	59.11	36.54	35.51	**61.19**	58.07	47.11±13.43
SimCLR	52.86	30.99	58.76	36.19	37.93	58.76	60.15	47.95±12.47
MoCo	55.99	29.25	55.99	38.28	36.89	52.51	61.19	47.16±12.16
ContraWR	54.78	**33.45**	55.13	36.78	37.78	56.67	61.15	47.96±11.45
SSCL	53.12	29.45	55.43	34.56	36.21	57.87	62.21	46.98±13.15
Ou2022	52.86	28.56	55.12	31.34	35.12	54.12	62.12	45.61±13.49
MAML	52.41	27.87	53.76	33.81	34.21	53.71	54.17	44.28±11.71
EOA	55.79	31.98	55.93	35.76	38.12	37.21	61.62	45.20±12.07
Song2022	53.54	30.46	57.21	38.28	35.27	55.41	60.31	47.21±12.12
SWPC	**57.64**	31.25	**59.38**	**38.54**	**39.58**	60.76	**63.89**	**50.15**±13.21

TABLE IV
ACCs ON MI3 IN WITHIN-SUBJECT CLASSIFICATION

Approach	Subject													Average
	1	2	3	4	5	6	7	8	9	10	11	12	13	
CEC	51.1	53.6	45.2	48.9	47.1	51.3	58.1	**56.9**	78.1	57.8	57.2	**56.1**	43.1	54.2±8.7
JTS	52.1	57.8	53.1	57.8	53.4	54.2	61.2	56.3	91.2	52.1	**63.6**	54.7	52.3	58.4±10.4
BYOL	54.0	69.0	65.6	60.6	60.6	72.3	69.0	49.0	99.0	**64.0**	57.3	54.0	50.6	63.5±12.9
SimCLR	51.3	71.3	64.7	74.7	**64.7**	**76.3**	64.7	49.7	93.0	61.3	49.7	41.3	53.0	62.7±14.0
MoCo	55.0	66.7	**70.0**	68.3	55.0	58.3	**73.3**	50.0	93.3	55.0	56.7	50.0	51.7	61.8±12.3
ContraWR	52.5	64.1	66.3	72.2	58.1	66.2	71.1	51.3	98.0	58.3	61.1	53.7	51.2	63.4±12.6
SSCL	51.2	61.1	67.1	70.2	52.3	64.5	67.9	52.1	**99.1**	55.1	59.1	52.3	48.5	61.6±13.3
Ou2022	49.2	59.1	67.3	71.1	54.3	63.1	70.1	51.3	98.0	56.2	57.8	54.3	53.2	61.9±12.9
MAML	52.8	62.9	60.4	74.2	58.2	70.3	68.5	49.4	98.0	58.4	52.6	49.5	51.5	62.1±13.4
EOA	52.7	68.5	61.2	69.2	57.3	71.7	65.3	48.1	98.2	56.9	52.6	47.8	50.2	61.5±13.7
Song2022	**57.8**	67.2	63.9	71.7	55.5	68.4	64.3	48.6	98.0	61.8	55.2	50.2	52.1	62.7±12.8
SWPC	55.0	**71.7**	61.7	**76.7**	60.0	73.3	71.7	50.0	98.3	63.3	56.7	51.7	**55.0**	**65.0**±13.2

TABLE V
ACCs ON MI4 IN WITHIN-SUBJECT CLASSIFICATION

Approach	Subject							Average
	1	2	3	4	5	6	7	
CEC	62.31	89.23	71.25	72.15	63.56	76.23	76.56	73.04±9.07
JTS	73.21	94.34	76.28	75.45	67.56	82.31	85.43	79.23±8.87
BYOL	75.62	97.19	78.44	79.69	77.81	87.81	86.88	83.35±7.66
SimCLR	72.51	97.52	71.88	80.62	77.81	85.31	86.25	81.70±8.96
MoCo	**77.81**	96.88	75.94	83.44	72.81	88.12	85.31	82.90±8.22
ContraWR	72.12	96.48	**81.41**	78.34	73.12	83.12	83.65	81.18±8.16
SSCL	69.43	91.67	78.45	73.43	75.02	78.43	81.43	78.27±7.09
Ou2022	68.76	95.46	78.34	75.87	73.56	76.12	82.75	78.69±8.54
MAML	67.31	93.71	78.42	80.61	76.42	85.32	84.02	80.83±8.21
EOA	72.76	96.32	75.64	79.21	77.37	84.28	85.21	81.54±7.90
Song2022	75.82	97.61	76.32	76.32	75.42	85.65	86.73	81.84±8.54
SWPC	71.25	**97.81**	80.01	**84.69**	**79.38**	**89.06**	**87.81**	**84.29**±8.47

Only two classes (right-hand and both feet) were used. Subject 10 was removed, as his/her results were close to random.

4) MI4 was the 004-2014 dataset [13] recorded from 9 subjects. EEG signals were sampled at 250Hz. Only two classes (left-hand and right-hand) were used. We deleted

TABLE VI
ACCs ON MI1 IN CROSS-SUBJECT CLASSIFICATION

Approach	Subject							Average
	1	2	3	4	5	6	7	
CEC	66.73	45.12	58.23	53.23	53.12	72.13	48.19	56.68±9.77
JTS	75.43	48.21	64.32	61.78	57.24	79.25	57.14	63.34±10.86
BYOL	73.61	51.39	70.83	60.42	56.94	85.42	56.25	64.98±12.06
SimCLR	77.78	50.69	72.92	63.89	54.86	82.64	49.31	64.58±13.48
MoCo	81.94	50.01	59.03	**66.67**	59.03	85.42	60.42	66.07±13.01
ContraWR	**82.54**	48.78	66.54	61.45	59.13	81.12	57.78	65.33±12.46
SSCL	74.34	49.45	69.87	58.78	58.87	79.32	54.17	63.54±11.08
Ou2022	74.91	48.76	67.43	56.91	58.76	75.54	51.12	61.92±10.87
MAML	76.51	48.93	71.68	59.32	58.32	81.83	57.32	64.84±11.93
EOA	77.81	51.82	72.01	57.71	60.63	83.62	56.32	65.71±12.10
Song2022	76.27	50.86	74.32	60.91	57.36	83.91	57.84	65.92±12.19
SWPC	80.56	**52.08**	**75.69**	62.51	**61.11**	**86.81**	**59.03**	**68.26**±12.79

TABLE VII
ACCs ON MI2 IN CROSS-SUBJECT CLASSIFICATION

Approach	Subject							Average
	1	2	3	4	5	6	7	
CEC	42.17	22.56	36.75	37.98	32.87	38.67	41.25	36.04±6.68
JTS	51.87	25.37	41.76	**45.67**	35.43	43.26	45.23	41.23±8.55
BYOL	52.47	25.74	**50.04**	39.62	37.89	**52.12**	42.41	42.90±9.64
SimCLR	44.44	**29.17**	45.83	27.78	28.82	31.61	32.29	34.28±7.59
MoCo	41.67	23.61	43.41	38.89	34.03	48.61	39.24	38.49±7.95
ContraWR	52.23	25.54	45.21	38.32	38.76	45.87	45.12	41.58±8.50
SSCL	49.87	23.51	43.93	35.56	38.78	41.32	43.87	39.55±8.38
Ou2022	48.76	24.28	43.87	31.53	36.65	38.87	42.46	38.06±8.19
MAML	46.92	25.98	40.92	37.81	39.21	44.63	43.51	39.85±6.88
EOA	53.85	25.46	45.18	38.62	37.91	42.81	44.95	41.25±8.73
Song2022	55.91	24.74	46.32	40.83	38.62	48.65	46.94	43.14±9.85
SWPC	**56.25**	27.08	48.96	42.01	**41.67**	51.04	**48.26**	**45.04**±9.40

TABLE VIII
ACCs ON MI3 IN CROSS-SUBJECT CLASSIFICATION

Approach	Subject													Average
	1	2	3	4	5	6	7	8	9	10	11	12	13	
CEC	**55.6**	57.3	62.1	62.3	51.3	58.2	57.4	71.3	74.3	51.2	42.1	47.8	51.2	57.1±9.1
JTS	51.2	68.2	64.3	67.6	58.2	71.2	63.5	71.2	75.5	53.4	45.7	52.6	50.2	61.2±9.6
BYOL	44.7	69.7	69.7	66.3	66.3	73.0	68.0	79.7	84.7	51.3	44.7	56.3	48.0	63.3±13.1
SimCLR	46.7	70.0	68.3	**73.3**	63.3	**76.7**	58.3	78.3	85.0	**61.7**	50.0	55.0	48.3	64.2±12.3
MoCo	51.7	**76.7**	66.7	**73.3**	53.3	65.0	56.7	65.0	**93.3**	50.0	50.0	**60.0**	53.3	62.7±12.7
ContraWR	53.9	70.3	64.9	71.1	61.7	75.4	67.8	76.5	87.2	52.8	51.7	52.3	48.3	64.1±11.9
SSCL	52.3	67.1	62.3	67.3	64.1	71.3	65.7	71.3	85.4	50.6	52.1	51.9	**54.1**	62.7±10.3
Ou2022	53.1	65.4	61.8	65.4	67.2	67.8	62.8	70.3	81.5	51.8	54.4	53.7	52.8	62.2±8.8
MAML	52.7	63.2	65.9	65.3	63.6	69.3	65.3	78.4	85.3	54.2	48.2	52.3	48.3	62.5±11.3
EOA	53.4	65.3	68.2	68.5	65.4	73.2	68.4	79.2	82.5	52.3	44.6	51.4	50.4	63.3±11.8
Song2022	52.3	68.4	67.3	70.2	61.4	72.9	**72.5**	81.9	83.7	51.2	**56.2**	50.2	51.7	64.6±11.7
SWPC	50.0	73.3	**70.0**	71.7	**68.3**	75.0	70.0	**83.3**	88.3	53.3	46.7	56.7	51.7	**66.2**±13.2

the results of the 2nd and 3rd subjects for the same reason.

Note that Subjects 5 and 6 in MI1, Subject 10 in MI3, and Subjects 2 and 3 in MI4, were removed because their results were close to random.

All EEG signals were bandpass-filtered between 8Hz and 30Hz, and then notch-filtered at 50Hz.

B. Performance Evaluation

The classification accuracy (ACC) was used as the performance metric. The specific computation details are illustrated in Fig. 8, where the yellow bar indicates the true MI period:

1) When all sliding windows during the true MI period have $\bar{p}_i \geq \tau$, $\hat{p}_i$ corresponding to the last sliding window with $\bar{p}_i \geq \tau$ is used to evaluate the classification accuracy.

2) When the true MI period is broken into two or more intervals, during each of which all $\bar{p}_i \geq \tau$, $\hat{p}_i$ corresponding to the last sliding window with $\bar{p}_i \geq \tau$ in the last interval is used to evaluate the classification accuracy.

3) When no sliding window in the MI period has $\bar{p}_i \geq \tau$, the classification is 'wrong'.

C. Algorithms

SWPC used the EEGNet backbone. Supervised learning used learning rate 0.0005 and early stopping with patience 30. SSL used learning rate 0.00005 and 40 training epochs.

SWPC was compared with the following 11 approaches:

1) Continuous EEG classification (CEC) [14], which uses CSP and thresholding to identify MIs in EEGs.

TABLE IX
ACCs on MI4 in Cross-Subject Classification

Approach	Subject							Average
	1	2	3	4	5	6	7	
CEC	60.15	81.23	75.45	72.34	68.43	76.54	78.23	73.20±7.07
JTS	67.32	86.24	75.34	78.12	75.23	**86.12**	81.23	78.51±6.72
BYOL	**77.81**	87.19	72.51	**82.81**	64.06	84.69	72.50	77.37±8.21
SimCLR	62.81	90.62	67.19	62.81	69.06	78.44	78.75	72.81±10.25
MoCo	73.75	90.62	74.38	80.62	**76.25**	79.69	50.62	75.13±12.22
ContraWR	72.45	84.98	74.98	78.42	72.98	82.76	81.65	78.32±4.99
SSCL	67.53	85.43	71.87	74.91	67.43	81.44	77.65	75.18±6.83
Ou2022	66.37	81.91	73.42	80.41	71.25	80.41	73.62	75.34±5.75
MAML	68.32	85.43	72.91	73.98	73.81	72.96	83.51	75.85±6.22
EOA	72.51	83.81	75.73	80.61	69.61	78.71	82.61	77.66±5.28
Song2022	76.26	87.53	73.61	82.71	73.27	73.81	80.51	78.24±5.49
SWPC	70.01	**91.88**	**76.88**	80.94	75.94	80.62	**85.01**	**80.18±6.99**

TABLE X
MI Identification Accuracies of the Prescreening Module

	Scenario	MI1	MI2	MI3	MI4	Average
Cross-subject	Without SSL in Prescreening	90.74	85.80	89.14	94.62	90.08
	With SSL in Prescreening	**92.69**	**86.92**	**91.89**	**94.17**	**91.41**
Within-subject	Without SSL in Prescreening	91.52	87.42	91.26	95.06	91.35
	With SSL in Prescreening	**93.74**	**89.19**	**93.21**	**96.19**	**93.08**

TABLE XI
ACCs of the Classification Module, When There Are Triggers

	Scenario	MI1	MI2	MI3	MI4	Average
Cross-subject	Without SSL in Classification	72.21	50.41	72.07	82.77	69.37
	With SSL in Classification	**75.39**	**52.95**	**73.81**	**84.56**	**71.67**
Within-subject	Without SSL in Classification	77.56	55.53	70.48	86.51	72.52
	With SSL in Classification	**79.83**	**56.79**	**72.14**	**88.34**	**74.28**

TABLE XII
ADJUSTED *p*-VALUES OF PAIRED *t*-TESTS BETWEEN SWPC AND OTHER APPROACHES IN WITHIN-SUBJECT CLASSIFICATION

SWPC vs.	MI1	MI2	MI3	MI4
CEC	0.0140*	0.0018*	0.0172*	0.0041*
JST	0.0140*	0.0015*	0.0126*	0.0140*
BYOL	0.0538	0.0721	0.0140*	0.0467*
SimCLR	0.0647	0.0361*	0.0752	0.0132*
MoCo	0.0291*	0.0227*	0.0698	0.0231*
ContraWR	0.0783	0.0658	0.0140*	0.0729
SSCL	0.0140*	0.0140*	0.0545	0.0140*
Ou2022	0.0140*	0.0140*	0.0140*	0.0158*
MAML	0.0778	0.0227*	0.0782	0.0136*
EOA	0.0432*	0.0285*	0.0589	0.0267*
Song2022	0.0231*	0.0782	0.0140*	0.0159*

TABLE XIII
ADJUSTED *p*-VALUES OF PAIRED *t*-TESTS BETWEEN SWPC AND OTHER APPROACHES IN CROSS-SUBJECT CLASSIFICATION

SWPC vs.	MI1	MI2	MI3	MI4
CEC	0.0037*	0.0067*	0.0013*	0.0027*
JST	0.0167*	0.0196*	0.0085*	0.0956
BYOL	0.0187*	0.0063*	0.0018*	0.0073*
SimCLR	0.0213*	0.0196*	0.1088	0.0027*
MoCo	0.0752	0.0027*	0.0196*	0.0031*
ContraWR	0.0292*	0.0596	0.0178*	0.0386*
SSCL	0.0358*	0.0196*	0.0478*	0.0196*
Ou2022	0.0169*	0.0196*	0.0523	0.0196*
MAML	0.0782	0.0596	0.0178*	0.0386*
EOA	0.0358*	0.0196*	0.0478*	0.0196*
Song2022	0.0171*	0.0782	0.0523	0.0782

2) Joint training scheme (JTS) [15], which combines the transitional imagery data (between the resting state and MI) with the resting state to train a binary classifier.
3) Bootstrap Your Own Latent (BYOL)[2] [16].
4) Simple framework for Contrastive Learning of Representations (SimCLR)[3] [17].
5) Momentum Contrast (MoCo)[4] [18].
6) ContraWR[5] [19].

7) Self-supervised contrastive learning (SSCL) [20], which was proposed for cross-session MI classification.
8) Ou2022 [21], an SSL approach for MI classification.
9) Model-agnostic meta-learning (MAML) [22], which learns a good initialization for fast adaptation.
10) Ensemble of averages (EOA) [23], which trains an ensemble of independent moving average models.
11) Song2022 [24], an event-related desynchronization detection and false positive rejection algorithm based on the time-frequency characteristics of MI.

Note that CEC and Song2022 were proposed for asynchronous MI classification, so their original algorithms were

[2]https://github.com/deepmind/deepmind-research/tree/master/byol
[3]https://github.com/google-research/simclr
[4]https://github.com/facebookresearch/moco
[5]https://github.com/ycq091044/ContraWR

TABLE XIV
ABLATION STUDY RESULTS IN WITHIN-SUBJECT CLASSIFICATION

SSL in Prescreening	SSL in Classification	Averaging	MI1	MI2	MI3	MI4	Average
×	×	×	68.36	44.01	58.52	81.32	63.05
√	×	×	69.45	46.12	60.04	82.15	64.44
×	√	×	70.93	48.43	64.03	82.34	66.43
×	×	√	69.76	47.17	59.13	83.43	64.87
√	√	×	71.24	49.31	64.22	83.39	67.04
√	×	√	70.98	48.54	64.32	82.24	66.52
×	√	√	71.87	49.41	64.15	83.91	67.34
√	√	√	**72.02**	**50.15**	**65.00**	**84.29**	**67.87**

TABLE XV
ABLATION STUDY RESULTS IN CROSS-SUBJECT CLASSIFICATION

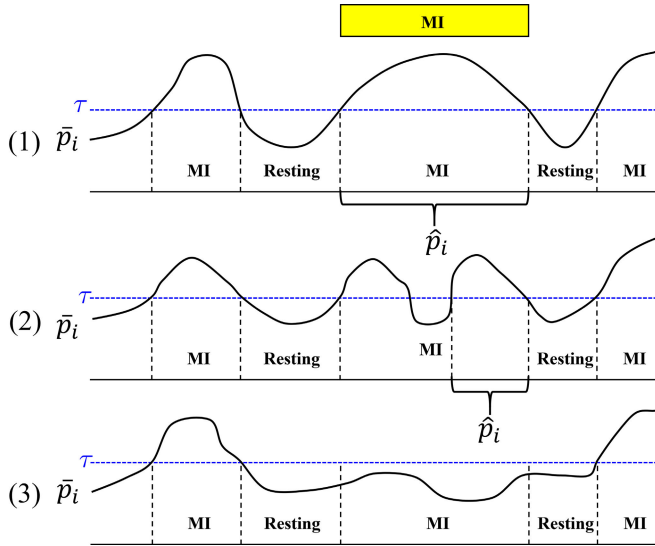| SSL in Prescreening	SSL in Classification	Averaging	MI1	MI2	MI3	MI4	Average
×	×	×	63.52	41.52	62.04	78.56	61.41
√	×	×	65.65	42.53	64.36	79.32	62.97
×	√	×	66.81	43.47	65.17	79.54	63.75
×	×	√	64.25	42.18	63.76	78.47	62.17
√	√	×	67.16	44.27	65.32	79.57	64.08
√	×	√	65.51	43.25	64.24	78.32	62.83
×	√	√	67.37	44.24	64.56	79.56	63.93
√	√	√	**68.26**	**45.04**	**66.20**	**80.18**	**64.88**

Fig. 8. Illustration of computing the classification accuracy in testing.

implemented. The other 9 algorithms cannot be directly used for asynchronous BCIs, so they were embedded into SWPC. More specifically, BYOL, SimCLR, MoCo, ContraWR, SSCL and Ou2022 were used to replace the SSL part of SWPC (the supervised learning part was kept), MAML and EOA were used to replace the supervised training part of SWPC (the SSL part was removed), and JTS was only used in supervised training of the prescreening module.

Both within-subject and cross-subject classifications were performed. For within-subject experiments, Session 1 was used for training, and Session 2 (with all triggers removed) of the same subject for testing. For cross-subject experiments, Session 2 of a subject was used for testing, and Session 1 from all other subjects were combined for training. 40% of the training data were reserved as the validation set for determining the optimal hyperparameters, which were then applied to all training data to re-train the model. Except for CEC, which has no randomness, all other algorithms were repeated five times, and the average is reported.

D. Results

The ACCs and standard deviations (across different subjects) of different approaches on the four MI datasets in within-subject classification are shown in Tables II-V, respectively. The ACCs and standard deviations in cross-subject classification are shown in Tables VI-IX, respectively. The best results are marked in bold. Clearly, our proposed SWPC achieved the best average results on all four datasets in both within-subject and cross-subject classifications.

We also studied the effectiveness of SSL to the prescreening module and the classification module. Table X shows the MI identification accuracies of the prescreening module, with and without SSL. Table XI shows the ACCs of the classification module, assuming there are triggers. Clearly, SSL was always beneficial to both the prescreening module and the classification module.

Paired t-tests were performed to evaluate whether the performance improvements of SWPC over others were statistically significant in within-subject and cross-subject classifications. The results are shown in Tables XII and XIII, respectively, where p-values smaller than 0.05 are marked with asterisks. Most of the performance improvements were statistically significant; particularly, in cross-subject classification, SWPC statistically significantly outperformed each algorithm on at least three datasets.

E. Ablation Study

Ablation studies were performed to evaluate if SSL in the prescreening module and the classification module, and the final averaging, are truly necessary and beneficial. Within-subject and cross-subject classification results are shown in
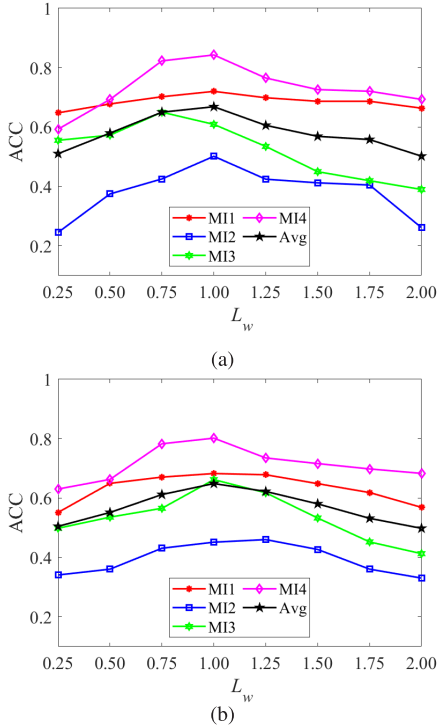
Fig. 9. Change of the classification accuracy w.r.t. the time window length L_w in (a) within-subject classification; and, (b) cross-subject classification.

Fig. 10. Change of the classification accuracy w.r.t. the prescreening threshold τ in (a) within-subject classification; and, (b) cross-subject classification.

TABLE XVI
ONLINE AND OFFLINE WITHIN-SUBJECT CLASSIFICATION ACCURACIES

Approach	Scenario	
	Online	Offline
BYOL	65.76	66.63
SimCLR	65.53	67.83
MoCo	65.60	66.54
ContraWR	65.58	67.83
SSCL	63.57	64.87
Ou2022	63.62	65.72
SWPC	**67.87**	**69.10**

TABLE XVII
ONLINE AND OFFLINE CROSS-SUBJECT CLASSIFICATION ACCURACIES

Approach	Scenario	
	Online	Offline
BYOL	62.13	64.64
SimCLR	58.97	62.83
MoCo	60.60	63.54
ContraWR	62.34	65.43
SSCL	60.25	62.76
Ou2022	59.37	63.21
SWPC	**64.88**	**67.82**

Tables XIV and XV, respectively. Clearly, all three components were essential to the superior performance of SWPC.

F. Parameter Sensitivity Analysis

This subsection evaluates the sensitivity of SWPC to the time window length L_w and the prescreening threshold τ. The results are shown in Figs. 9 and 10, respectively. $L_w = 1$ and $\tau = 0.2$ seem to achieve the overall best performance on all datasets.

G. Offline Classification

All previous subsections considered online classification, i.e., the test data are available on-the-fly. This subsection further considers offline classification, where all test EEG data are available.

In offline classification, SSL on the test set $\mathcal{D}_t = \{X_i^t\}_{i=1}^{n_t}$ (instead of on the training set $\mathcal{D}_s$ in online classification) may be used to improve the performance. Specifically, we first conducted SSL in Section II-C on $\mathcal{D}_t$, and then SSL in
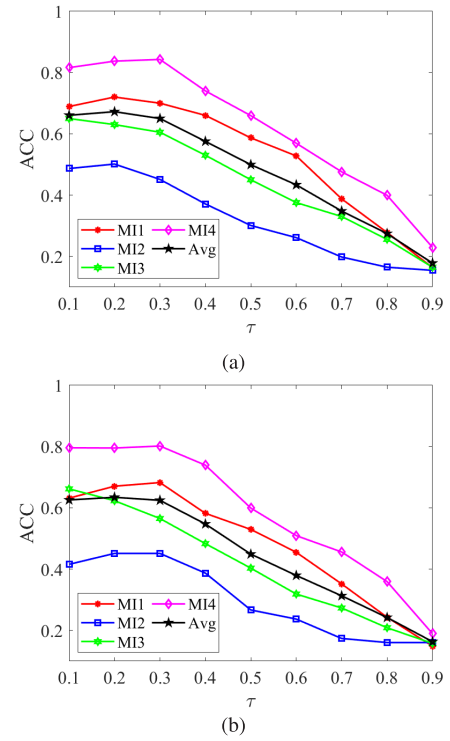
Section II-D on $\hat{\mathcal{D}}_t$, which consisted of EEG trials predicted as MI in $\mathcal{D}_t$. Tables XVI and XVII show the average results on the four datasets in within-subject and cross-subject classification, respectively. Offline classification accuracies were higher than their online counterparts for all approaches in both scenarios, because SSL on the test data themselves extracted more tailored features.

IV. CONCLUSION AND FUTURE RESEARCH

Asynchronous MI-based BCIs aim to detect the user's MI without explicit triggers. They are challenging to implement, because the algorithm needs to first distinguish between resting-states and MI trials, and then classify the MI trials into the correct task, all without any triggers. This paper has proposed SWPC for MI-based asynchronous BCIs, which consists of two modules: a prescreening module to screen MI trials from the resting-state, and a classification module for MI classification. Both modules are trained with supervised learning followed by SSL. Within-subject and cross-subject asynchronous MI classification on four different EEG datasets

validated the effectiveness of SWPC, particularly, SSL to refine the feature extractors.

Our future research directions include:

1) *Transfer learning*: Transfer learning can further mitigate cross-subject and cross-session data discrepancies. For asynchronous BCIs, data alignment [25], source-free domain adaptation [26], and domain generalization [27] approaches may be used to further improve performance and protect user privacy.

2) *Test-time adaptation*: Test-time adaptation updates the classifier using online unlabeled data to improve its performance. Our recent work [28] has demonstrated its promising performance in synchronous MI-based BCIs, but how to apply it to asynchronous BCIs requires further investigation.

3) *More BCI paradigms*: Only MI was considered in this paper. It is interesting to study if SWPC can be extended to other classical BCI paradigms, e.g., event-related potential and steady-state visual evoked potential.

REFERENCES

[1] B. J. Lance, S. E. Kerick, A. J. Ries, K. S. Oie, and K. McDowell, "Brain–computer interface technologies in the coming decades," *Proc. IEEE*, vol. 100, no. Special Centennial Issue, pp. 1585–1599, May 2012.

[2] D. Wu, X. Jiang, and R. Peng, "Transfer learning for motor imagery based brain–computer interfaces: A tutorial," *Neural Netw.*, vol. 153, pp. 235–253, Sep. 2022.

[3] D. Wu, B.-L. Lu, B. Hu, and Z. Zeng, "Affective brain–computer interfaces (aBCIs): A tutorial," *Proc. IEEE*, vol. 111, no. 10, pp. 1314–1332, Oct. 2023.

[4] 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.

[5] K. K. Ang et al., "A clinical study of motor imagery-based brain-computer interface for upper limb robotic rehabilitation," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 5981–5984.

[6] X. Zhang, L. Yao, Q. Z. Sheng, S. S. Kanhere, T. Gu, and D. Zhang, "Converting your thoughts to texts: Enabling brain typing via deep feature learning of EEG signals," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Athens, Greece, Mar. 2018, pp. 1–10.

[7] A. Palumbo, V. Gramigna, B. Calabrese, and N. Ielpo, "Motor-imagery EEG-based BCIs in wheelchair movement and control: A systematic literature review," *Sensors*, vol. 21, no. 18, p. 6285, Sep. 2021.

[8] S. G. Mason and G. E. Birch, "A brain-controlled switch for asynchronous control applications," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 10, pp. 1297–1307, Oct. 2000.

[9] T. Sugiura, N. Goto, and A. Hayashi, "A discriminative model corresponding to hierarchical HMMs," in *Proc. 14th Int. Conf. Intell. Data Eng. Automated Learn.*, Birmingham, U.K., Dec. 2007, pp. 375–384.

[10] J. F. Delgado Saa and M. Çetin, "Discriminative methods for classification of asynchronous imaginary motor tasks from EEG data," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 5, pp. 716–724, Sep. 2013.

[11] M. Tangermann et al., "Review of the BCI competition IV," *Frontiers Neurosci.*, vol. 45, no. 2, pp. 55–62, 2012.

[12] D. Steyrl, R. Scherer, J. Faller, and G. R. Müller-Putz, "Random forests in non-invasive sensorimotor rhythm brain-computer interfaces: A practical and convenient non-linear classifier," *Biomed. Eng./Biomedizinische Technik*, vol. 61, no. 1, pp. 77–86, Feb. 2016.

[13] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain–computer communication: Motivation, aim, and impact of exploring a virtual apartment," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 4, pp. 473–482, Dec. 2007.

[14] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 2, pp. 258–265, Jun. 2004.

[15] P. Cheng, P. Autthasan, B. Pijarana, E. Chuangsuwanich, and T. Wilaiprasitporn, "Towards asynchronous motor imagery-based brain-computer interfaces: A joint training scheme using deep learning," in *Proc. TENCON IEEE Region 10 Conf.*, Oct. 2018, pp. 1994–1998.

[16] J.-B. Grill et al., "Bootstrap your own latent—A new approach to self-supervised learning," in *Proc. Conf. Neural Inf. Process. Syst.*, Dec. 2020, pp. 21271–21284.

[17] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *Proc. Int. Conf. Mach. Learn.*, 2020, pp. 1597–1607.

[18] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 9726–9735.

[19] C. Yang, C. Xiao, M. B. Westover, and J. Sun, "Self-supervised electroencephalogram representation learning for automatic sleep staging: Model development and evaluation study," *JMIR AI*, vol. 2, Jul. 2023, Art. no. e46769.

[20] T. Lotey, P. Keserwani, G. Wasnik, and P. P. Roy, "Cross-session motor imagery EEG classification using self-supervised contrastive learning," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 975–981.

[21] Y. Ou, S. Sun, H. Gan, R. Zhou, and Z. Yang, "An improved self-supervised learning for EEG classification," *Math. Biosci. Eng.*, vol. 19, no. 7, pp. 6907–6922, 2022.

[22] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. Int. conf. Mach. Learn.*, Sydney, NSW, Australia, Aug. 2017, pp. 1126–1135.

[23] D. Arpit, H. Wang, Y. Zhou, and C. Xiong, "Ensemble of averages: Improving model selection and boosting performance in domain generalization," in *Proc. Adv. Neural Inf. Process. Syst.*, New Orleans, LA, USA, Nov. 2022, pp. 8265–8277.

[24] M. Song, H. Jeong, J. Kim, S.-H. Jang, and J. Kim, "An EEG-based asynchronous MI-BCI system to reduce false positives with a small number of channels for neurorehabilitation: A pilot study," *Frontiers Neurorobotics*, vol. 16, pp. 215–230, Sep. 2022.

[25] H. He and D. Wu, "Transfer learning for brain–computer interfaces: A Euclidean space data alignment approach," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 2, pp. 399–410, Feb. 2020.

[26] K. Xia, L. Deng, W. Duch, and D. Wu, "Privacy-preserving domain adaptation for motor imagery-based brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 11, pp. 3365–3376, Nov. 2022.

[27] S. Li, H. Wu, L. Ding, and D. Wu, "Meta-learning for fast and privacy-preserving source knowledge transfer of EEG-based BCIs," *IEEE Comput. Intell. Mag.*, vol. 17, no. 4, pp. 16–26, Nov. 2022.

[28] S. Li, Z. Wang, H. Luo, L. Ding, and D. Wu, "T-TIME: Test-time information maximization ensemble for plug-and-play BCIs," *IEEE Trans. Biomed. Eng.*, vol. 71, no. 2, pp. 423–432, Feb. 2024.