Subject-Independent Brain-Computer Interfaces with Open-Set Subject Recognition

Dong-Kyun Han
Dept. Brain and Cognitive Engineering
Korea University
Seoul, Republic of Korea
dk_han@korea.ac.kr

Geun-Deok Jang
Dept. Artificial Intelligence
Korea University
Seoul, Republic of Korea
gd_jang@korea.ac.kr

Dong-Young Kim
Dept. Artificial Intelligence
Korea University
Seoul, Republic of Korea
dy_kim@korea.ac.kr

Abstract—A brain-computer interface (BCI) can't be effectively used since electroencephalography (EEG) varies between and within subjects. BCI systems require calibration steps to adjust the model to subject-specific data. It is widely acknowledged that this is a major obstacle to the development of BCIs. To address this issue, previous studies have trained a generalized model by removing the subjects' information. In contrast, in this work, we introduce a style information encoder as an auxiliary task that classifies various source domains and recognizes open-set domains. Open-set recognition method was used as an auxiliary task to learn subject-related style information from the source subjects, while at the same time helping the shared feature extractor map features in an unseen target. This paper compares various OSR methods within an open-set subject recognition (OSSR) framework. As a result of our experiments, we found that the OSSR auxiliary network that encodes domain information improves generalization performance.

Keywords—*Brain–computer interface; electroencephalography; motor imagery; domain generalization; open-set recognition*

I. INTRODUCTION

The brain-computer interface (BCI) interprets the intention of the user to communicate with external devices by analyzing brain signals [1]–[3]. Among the various methods for measuring brain signals [4]–[7], a well-established and widely used brain signal is electroencephalography (EEG), which is non-invasive and has a high temporal resolution [8], [9]. The following paradigms are commonly used for EEG-based BCI: motor imagery (MI) [10]–[13], event-related potential (ERP) [14]–[17], and steady-state visual potential (SSVEP) [18].

It is challenging to analyze EEGs because they vary over time and between subjects due to psychological or physiolog-

ical changes. A disadvantage of this intra-subject variability is that it requires subject-specific calibration each time a new user uses the BCI [19]. In addition to collecting subject-specific data and tuning the model, calibration takes approximately 20-30 minutes [20]–[22]. BCI requires a reduction or elimination of this calibration procedure for practical application.

With the rise of deep learning, from a domain shift perspective, many previous studies have proposed transfer learning-based approaches [23]–[26]. In this context, the problem of training generalized BCI models for an unknown subject, known as subject-independent BCI, falls under the domain generalization (DG) problem.

Previous subject-independent studies for DG purposes, either explicitly or implicitly, have been proposed in the direction of eliminating or ignoring subject (domain) information. It is possible, however, that in a transfer learning process, data from other subjects may have negative effects. In light of this, it may be beneficial to learn subject-specific features simultaneously, i.e., to give the network the ability to distinguish which individuals a sample belongs to.

In [27], this issue was addressed by jointly training an auxiliary network that performs an open-set recognition (OSR) task to learn subject-specific style features and to impart invariance between instances of the same subject. OSR aims to provide a system capable of identifying known and unknown classes for real-world scenarios in which unknown classes might be encountered. The main objective is to lower both the closed-set classification risk associated with labeled known data as well as the open space risk associated with unknown data at the same time [28]. Despite the fact that existing closed-set methods (i.e. softmax) are good at distinguishing classes, their ability to distinguish between known and unknown classes is limited. Therefore, several methods [29], [30] employ prototype learning as a means of forcing training features to resemble the corresponding prototypes, making the distinction between known and unknown easier.

In this study, we validated the effectiveness of the OSR task for learning subject-specific style features in a prototype-based domain generalization framework based on [27]. Various OSR methods, which we refer to as open-set subject recogni-

tion (OSSR), were compared against various combinations of subject pools.

II. METHODS

OSSR task utilizes the subject labels to impose cross-instance style (subject-specific information) invariance and to learn subject discriminative features, rather than remove user information. A basic framework configuration consists of the following. Style encoder and semantic encoder follow convolutional neural network (CNN). CNN is employed to extract common features between the encoders. This style encoder serves as a tool for supporting the classification of subjects by enabling the feature extractor to recognize them. Using the OSR method as an auxiliary task to classify subjects reduces the open space risk of potentially unknown subjects and trains a more generalized model. We also use OSR methods to classify tasks based on semantic features that are derived from the semantic encoder. An overview of the OSSR framework is shown in Fig. 1. It should be noted that the OSSR framework's goal is not to classify the subjects accurately, but rather to use knowledge from the subjects in order to assist in task classification.

A. Prototype Learning for EEG Decoding

In this section, we describe the prototype learning methods we used in our experiments. The prototype learning used in our experiments was originally proposed for open-world problems and open-set recognition. In this study, we compare various combinations of OSSR frameworks using the convolutional prototype learning (CPL)-based OSR [29] used in [27] and the following prototype learning-based OSR methodologies reciprocal points learning (RPL) [31] and adversarial reciprocal points learning (ARPL) [32].

A prototype is an average or representative example of each class that expresses the characteristics of the entire instance of each class. In contrast to traditional CNNs, prototype learning does not use a softmax layer but instead learns prototypes based on a data set. Here, prototypes are learnable representations formed by one or more latent features. The following sections provide a detailed description of the three methodologies described above.

1) Generalized Convolutional Prototype Learning: Yang *et al.* [29] proposed the Generalized Convolutional Prototype Learning (GCPL). In GCPL prototypes are trained in conjunction with a feature extractor and instances are classified according to the most similar prototype. Additionally, a prototype loss (PL) is proposed as a regularization for enhancing the intra-class compactness of the representation.

In [27], they used hybrid loss with distance-based cross-entropy loss (DCE) loss and PL, just like GCPL. At this time, the GCPL loss was also used for the semantic encoder ($\mathcal{L}_{clf}$). Following that, we add an experiment with a framework in which the semantic encoder is trained with cross-entropy loss, i.e., the semantic task is a closed-set classification.

In this approach, the distance between the samples and the prototypes is considered the probability of that sample

Fig. 1. The overview of the proposed framework. Our framework consists of two encoders, one for style and one for semantic. In feature space, different colors denote different subjects, and different shapes denote different classes.

belonging to the class of that prototype. Let's assume there are C classes and each has one prototype. Considering the training sample (x, y), let M_i represents the prototype where $i \in (1, 2, ..., C)$. The prototypes $M = \{m_i | i = 1, ..., C\}$ are learned during training.

The probability that the sample x belongs to class k is related to the probability that the extracted feature belongs to the prototype m. The probability for the prototype m is measured by the corresponding distance and softmax as:

$$p(y = k | x, M) = \frac{e^{-\gamma(\|f(x) - m_i\|_2^2)}}{\sum_{k=1}^{C} e^{-\gamma(\|f(x) - m_k\|_2^2)}}, \quad (1)$$

where $f(x)$ is the CNN-based feature extractor and γ is a temperature parameter that controls the probability assignment hardness. DCE is the cross-entropy loss calculated using this probability. As a regularizer, PL is added as follows:

$$l_p = \| f(x) - m_i \|_2^2. \quad (2)$$

As a result, the combined loss function is as follows:

$$l_{gcpl} = l_{dce} + \beta l_p. \quad (3)$$

The weight of the PL is controlled by β.

2) Reciprocal Points Learning: The RPL [31] introduced a new concept, Reciprocal Point, which represents the extra-class space corresponding to each known class. As with GCPL, the learnable parameter reciprocal point M is calculated by using the following loss function.

$$p(y = k | x, M) = \frac{e^{\gamma(\|f(x) - m_i\|_2^2)}}{\sum_{k=1}^{C} e^{\gamma(\|f(x) - m_k\|_2^2)}}, \quad (4)$$

$$l_{rp} = -\log p(y = k | x, M). \quad (5)$$

RPL introduced a regularization term that limits the distance between the closed-set sample and reciprocal point to some extent.

$$l_o(x; M^k, R^k) = MSE(\| f(x) - m_k \|_2^2, R^k), \quad (6)$$

where R^k is the radius (learnable margin) initialized to 1. The combined loss function is as follows:

$$l_{rpl} = l_{rp} + \gamma l_o, \quad (7)$$

where γ is a hyperparameter to control the weight of reducing open space risk module.

3) Adversarial Reciprocal Points Learning: ARPL [32] is based on RPL, but the distance and constraint functions have been changed. The distance $d(f(x), M^k)$ between x and reciprocal point M^k can be obtained by combining the Euclidean distance d_e with the dot product d_d:

$$d(f(x), M^k) = d_e(f(x), M^k) - d_d(f(x), M^k). \quad (8)$$

The classification probability can be expressed as follows:

$$p(y = k|x, M) = \frac{e^{\gamma d(f(x), M^i)}}{\sum_{k=1}^{C} e^{\gamma d(f(x), M^k)}}. \quad (9)$$

In ARPL, the regularization term was changed as follows:

$$l_o(x; M^k, R^k) = MAX(d_e(f(x), M^k) - R, 0). \quad (10)$$

Margin R is also initialized to 1.

B. Open-Set Subject Recognition framework

As shown in Fig. 1, in order to separate subject information from class information, we employ two encoders referred to as style and semantic encoders. In our framework, each encoder can be trained by various types of Loss. We train the style and semantic encoder using the hybrid loss function as follows:

$$\mathcal{L} = \mathcal{L}_{clf} + \alpha \mathcal{L}_{ossr}, \quad (11)$$

where $\mathcal{L}_{clf}$ means the loss function calculated for semantic encoder, and $\mathcal{L}_{ossr}$ means the loss function calculated for style encoder. α controls the weight of the OSSR task.

III. EXPERIMENTS

A. Dataset and Data-Split

OpenBMI dataset [33]: The dataset includes 54 subjects and two motor imagery classes (left hand, right hand). Data for each subject consists of 4 sessions: offline and online sessions for two days. There are 100 trials in each session. Each trial consists of 4 seconds of EEG recorded at 62 channels, 1000Hz, and we downsampled to 250Hz for this experiment. We used Leave-one-subject-out cross-validation. Data from the source dataset was divided into two parts, 8:2, and used as training and validation data, respectively. For evaluation, only the fourth session of the test subject was considered [34]. We experimented on datasets consisting of four different sizes of subject numbers. There were 11 subjects, 21 subjects, 31 subjects, and 54 subjects. The runs contain different sets of subjects. However, for each method, the same sets are used. Five runs are conducted with 10 subjects, three runs with 20 subjects, two runs with 30 subjects, and one run with 53 subjects. The procedure of all training follows the data split configuration of [27].

TABLE I
MOTOR IMAGERY CLASSIFICATION PERFORMANCE (ACCURACY (%)).
RESULTS ARE AVERAGED AMONG RUNS.

Method	# subject			
	10	20	30	53
Baseline [27]	72.83 ($\pm$14.22)	80.65 ($\pm$13.02)	82.81 ($\pm$12.83)	84.98 ($\pm$12.18)
GCPL$_{clf}$ [27]	72.67 ($\pm$14.04)	80.30 ($\pm$12.45)	82.22 ($\pm$12.11)	84.80 ($\pm$11.84)
GCPL$_{clf}$+GCPL$_{ossr}$ [27]	**74.17** ($\pm$14.19)	**81.61** ($\pm$12.83)	**84.07** ($\pm$12.25)	85.22 ($\pm$12.24)
CE$_{clf}$+GCPL$_{ossr}$	73.67 ($\pm$12.78)	79.99 ($\pm$12.88)	82.97 ($\pm$12.11)	84.48 ($\pm$12.41)
RPL$_{clf}$+RPL$_{ossr}$	70.00 ($\pm$12.64)	77.39 ($\pm$13.99)	81.81 ($\pm$12.89)	84.06 ($\pm$12.44)
CE$_{clf}$+RPL$_{ossr}$	71.96 ($\pm$12.99)	79.73 ($\pm$13.19)	82.86 ($\pm$12.11)	84.57 ($\pm$12.17)
ARPL$_{clf}$+ARPL$_{ossr}$	73.34 ($\pm$13.73)	80.35 ($\pm$13.64)	82.39 ($\pm$12.89)	**85.33** ($\pm$11.73)
CE$_{clf}$+ARPL$_{ossr}$	72.71 ($\pm$12.71)	80.08 ($\pm$13.34)	82.96 ($\pm$12.16)	84.72 ($\pm$11.71)

B. Experimental Details

Following previous studies [27], [34], DeepConvNet [12] was used as a shared backbone feature extractor. In both style and semantic encoders, a fully connected layer (1400 $\times$ 2) was used. We set the case where $\mathcal{L}_{clf}$ is cross-entropy loss and $\mathcal{L}_{ossr}$ is not used as the baseline, and the case where l_{gcpl} is used for both $\mathcal{L}_{clf}$ and $\mathcal{L}_{ossr}$ is marked as [27]. β was set to 0.001 and the weight of $\mathcal{L}_{ossr}$ (α) was set to 0.1. We experimented with cases where cross-entropy or l_{rpl} or l_{arpl} was used for $\mathcal{L}_{clf}$ and l_{rpl} or l_{arpl} was used for $\mathcal{L}_{ossr}$. When l_{rpl} was used, γ was set to 0.001. In all cases, The weight of $\mathcal{L}_{ossr}$ (α) was set to 0.1. Adam was used as the optimizer and trained at a learning rate of 0.005, and a cosine annealing learning rate scheduler was also used.

IV. RESULTS AND DISCUSSION

Various OSSR results were compared with existing subject-independent methods. In total, we experimented with five different methods in subject-independent settings. The first is that the semantic encoder is trained with CE, while a style encoder is trained with GCPL and the second is the RPL is used for both the semantic encoder and style encoder. Thirdly, semantic encoders use CE, and style encoders use RPL. Fourth, ARPL is applied to both semantic and style encoders. Lastly, CE is applied to semantic encoders while ARPL is applied to style encoders. Comparing the results can be found in Table I.

CE$_{clf}$+ GCPL$_{ossr}$ achieves 73.67% average accuracy on 10 subjects. RPL$_{clf}$+ RPL$_{ossr}$ and CE$_{clf}$+ RPL$_{ossr}$ showed poor performance in 10 subjects, but good performance compared to the baseline in 30 and 53 and showed that reciprocal points can be used as a classifier. ARPL$_{clf}$+ ARPL$_{ossr}$ and CE$_{clf}$+ ARPL$_{ossr}$ showed similar performance to the baseline in all cases and was higher than RPLs in most cases. To summarize, ARPL$_{clf}$+ARPL$_{ossr}$ achieved the highest performance in 53 subjects, and GCPL$_{clf}$+ GCPL$_{ossr}$ showed the highest performance in the other cases. However, the proposed OSSR framework showed good performance compared to the baseline in the settings of 10 and 30 subjects and showed a slight difference even with 20 subjects. As mentioned in [27], performance increased as the number of subjects increased, but the gap between methods narrowed.

V. CONCLUSION

In this paper, we proposed an open-set subject recognition framework for subject-independent BCIs. In our framework, open-set recognition was used as an auxiliary task to encode subjects' information. The proposed framework has, to the best of our knowledge, achieved the best performance on the OpenBMI dataset using subject-independent settings. In the future, we intend to introduce a framework for subject-adaptive classification (domain adaptation). We also plan to apply the OSR framework, which classifies previously unseen classes, to real-world BCIs.

REFERENCES

[1] M. A. Lebedev and M. A. Nicolelis, "Brain–machine interfaces: past, present and future," *Trends Neurosci.*, vol. 29, no. 9, pp. 536–546, Sep. 2006.

[2] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[3] K.-T. Kim, H.-I. Suk, and S.-W. Lee, "Commanding a brain-controlled wheelchair using steady-state somatosensory evoked potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 3, pp. 654–665, Aug. 2016.

[4] N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: a review," *Front. Human Neurosci.*, vol. 9, p. 3, Jan. 2015.

[5] Y. Zhang, H. Zhang, X. Chen, S.-W. Lee, and D. Shen, "Hybrid high-order functional connectivity networks using resting-state functional mri for mild cognitive impairment diagnosis," *Scientific reports*, vol. 7, no. 1, pp. 1–15, 2017.

[6] K.-H. Thung, P.-T. Yap, E. Adeli, S.-W. Lee, D. Shen, A. D. N. Initiative *et al.*, "Conversion and time-to-conversion predictions of mild cognitive impairment using low-rank affinity pursuit denoising and matrix completion," *Medical image analysis*, vol. 45, pp. 68–82, 2018.

[7] M. Lee, B. Baird, O. Gosseries, J. O. Nieminen, M. Boly, B. R. Postle, G. Tononi, and S.-W. Lee, "Connectivity differences between consciousness and unconsciousness in non-rapid eye movement sleep: a TMS–EEG study," *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.

[8] Y. Chen, A. D. Atnafu, I. Schlattner, W. T. Weldtsadik, M.-C. Roh, H. J. Kim, S.-W. Lee, B. Blankertz, and S. Fazli, "A high-security EEG-based login system with RSVP stimuli and dry electrodes," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 12, pp. 2635–2647, Jun. 2016.

[9] J.-H. Jeong, N.-S. Kwak, C. Guan, and S.-W. Lee, "Decoding movement-related cortical potentials based on subject-dependent and section-wise spectral filtering," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 28, no. 3, pp. 687–698, Jan. 2020.

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

[11] J.-H. Cho, J.-H. Jeong, K.-H. Shim, D.-J. Kim, and S.-W. Lee, "Classification of hand motions within EEG signals for non-invasive BCI-based robot hand control," in *Conf. Proc. IEEE Int. Syst. Man Cybern. (SMC)*, Miyazaki, Japan, Oct. 2018, pp. 515–518.

[12] 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 EEG decoding and visualization," *Hum. Brain Mapp.*, vol. 38, no. 11, pp. 5391–5420, Aug. 2017.

[13] J.-H. Jeong, K.-H. Shim, D.-J. Kim, and S.-W. Lee, "Brain-controlled robotic arm system based on multi-directional CNN-BiLSTM network using EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1226–1238, Mar. 2020.

[14] R. Fazel, B. Z. Allison, C. Guger, E. W. Sellers, S. C. Kleih, and A. Kübler, "P300 brain computer interface: current challenges and emerging trends," *Front. Neuroeng*, vol. 5, p. 14, Jul. 2012.

[15] D.-O. Won, H.-J. Hwang, D.-M. Kim, K.-R. Müller, and S.-W. Lee, "Motion-based rapid serial visual presentation for gaze-independent brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 334–343, Aug. 2017.

[16] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components—a tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, May 2011.

[17] M.-H. Lee, J. Williamson, D.-O. Won, S. Fazli, and S.-W. Lee, "A high performance spelling system based on EEG-EOG signals with visual feedback," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 7, pp. 1443–1459, May 2018.

[18] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components," *J. Neural Eng*, vol. 2, no. 4, pp. 123–130, Oct. 2005.

[19] H.-I. Suk, S. Fazli, J. Mehnert, K.-R. Müller, and S.-W. Lee, "Predicting bci subject performance using probabilistic spatio-temporal filters," *PloS one*, vol. 9, no. 2, p. e87056, 2014.

[20] H.-I. Suk and S.-W. Lee, "Subject and class specific frequency bands selection for multiclass motor imagery classification," *Int. J. Imag. Syst. Tech.*, vol. 21, pp. 123–130, May 2011.

[21] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE. Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, 2010.

[22] M.-H. Lee, S. Fazli, J. Mehnert, and S.-W. Lee, "Subject-dependent classification for robust idle state detection using multi-modal neuroimaging and data-fusion techniques in BCI," *Pattern Recognit.*, vol. 48, pp. 2725–2737, Aug. 2015.

[23] K.-T. Kim, C. Guan, and S.-W. Lee, "A subject-transfer framework based on single-trial EMG analysis using convolutional neural network," *IEEE Trans. Neural Syst. Rehabil Eng.*, vol. 28, no. 1, pp. 94–103, 2019.

[24] C. E. Solórzano-Espíndola, E. Zamora, and H. Sossa, "Multi-subject classification of motor imagery EEG signals using transfer learning in neural networks," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2021, pp. 1006–1009.

[25] I. Dolzhikova, B. Abibullaev, R. Sameni, and A. Zollanvari, "An ensemble cnn for subject-independent classification of motor imagery-based EEG," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2021, pp. 319–324.

[26] O.-Y. Kwon, M.-H. Lee, C. Guan, and S.-W. Lee, "Subject-independent brain-computer interfaces based on deep convolutional neural networks," *IEEE Trans. Neural Networks Learn. Syst.*, Nov. 2019.

[27] S. Musellim, D.-K. Han, J.-H. Jeong, and S.-W. Lee, "Prototype-based domain generalization framework for subject-independent brain-computer interfaces," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Glasgow, Scotland, Jul. 2022, pp. 711–714.

[28] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1757–1772, 2013.

[29] H.-M. Yang, X.-Y. Zhang, F. Yin, and C.-L. Liu, "Robust classification with convolutional prototype learning," in *2018 Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3474–3482.

[30] H.-M. Yang, X.-Y. Zhang, F. Yin, Q. Yang, and C.-L. Liu, "Convolutional prototype network for open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2020.

[31] G. Chen, L. Qiao, Y. Shi, P. Peng, J. Li, T. Huang, S. Pu, and Y. Tian, "Learning open set network with discriminative reciprocal points," in *Proc. European Conference on Computer Vision (ECCV)*, 2020, pp. 507–522.

[32] G. Chen, P. Peng, X. Wang, and Y. Tian, "Adversarial reciprocal points learning for open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 8065–8081, 2022.

[33] M.-H. Lee, O.-Y. Kwon, Y.-J. Kim, H.-K. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-W. Lee, "EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy," *GigaScience*, vol. 8, no. 5, 2019.

[34] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network," *Neural. Netw.*, vol. 136, pp. 1–10, 2021.